GRUNDLINIEN
DES
WISSENSCHAFTLICHEN RECHNENS

VON

Dr. HEINRICH BRUNS,

PROFESSOR DER ASTRONOMIE AN DER UNIVERSITÄT ZU LEIPZIG.

LEIPZIG,
DRUCK UND VERLAG VON B. G. TEUBNER.
1903.

Vorwort.

Der Verfasser hatte bei den Übungen in seinem Seminar für „wissenschaftliches Rechnen“ schon vor längerer Zeit damit begonnen, den Teilnehmern die zur Vorbereitung erforderlichen mathematischen Entwickelungen autographiert in die Hand zu geben, um dadurch Zeit für die Beschäftigung mit besonderen Aufgaben zu gewinnen. Diese Aufzeichnungen werden hier in etwas erweiterter Gestalt der Öffentlichkeit übergeben, da es sich um Dinge handelt, für die es bisher an einer handlichen Zusammenstellung fehlte, und die überdies außerhalb des Kreises der berufsmäßigen Rechner keineswegs so bekannt sind, wie sie es bei ihrer erprobten Nützlichkeit verdienen.

Die Darstellung ist, da es sich in erster Linie um einen Leitfaden für den akademischen Unterricht handelt, auf die zum Verständnis unentbehrlichen Entwicklungen beschränkt: der Lehrer ist ohnehin genötigt, bei der Auswahl und Erläuterung der jedesmal zu stellenden Aufgaben auf die Vorbildung der Zuhörer Rücksicht zu nehmen.

Leipzig, 1903 März 9.

H. Bruns.

Inhalt.

Einleitung.

Seite

§ 1—3. Geschichtliches 1
§ 4. Gliederung der Rechentechnik 3
§ 5—6. Hilfsmittel; Maschinen 4
§ 7. Tafeln 7
§ 8. Gliederung der Darstellung 9

I. Differenzen und Summen.

§ 9—11. Das Differenzenschema; Erweiterung durch die Summenspalten; erzeugende Funktionen 11
§ 12. Das Umklappen des Schemas 14
§ 13—14. Das Schema einer Verbindung; Zwischengrößen; Fehlerschema 14
§ 15. Verhalten der Tafeldifferenzen 16
§ 16—17. Bezeichnung von *Gauß*; Haupt- und Zwischentabellen; Beispiel 17
§ 18—19. Analytische Darstellung der Differenzen; Bereich der Gültigkeit. 20
§ 20—21. Ganze Funktionen 23
§ 22. Kontrolle durch Differenzen 26
§ 23. Zweifelhafte Abrundung 27

II. Interpolation bei Tafeln.

§ 24. Aufgabe 29
§ 25. Formel von *Lagrange* 29
§ 26—29. Die Formel J; die Korrektion; die Konvergenz 31
§ 30—32. Die Formeln N, G, G', G'', S, B 37
§ 33. Zusammenhang mit der Formel von *Lagrange* 41
§ 34—39. Vergleichung und Benutzung von N, G, S, B 42
§ 40. Interpolation in die Mitte 50
§ 41—42. Interpolation bei Tabulierungen 51
§ 43—47. Wirkung der Abrundungsfehler 54

III. Numerische Differentiation.

§ 48. Grundformel 62
§ 49—51. Gebrauchsformeln 63

IV. Numerische Integration: Summenmethode.

§ 52. Allgemeine Bemerkungen 68
§ 53—55. Die Hauptformeln und die Bestimmung der Koeffizienten . . . 69
§ 56. Bestimmung der Konstanten 73
§ 57—58. Beispiel 75
§ 59. Allgemeine Bemerkungen 78
§ 60—62. Die Integration von Differentialgleichungen 80
§ 63. Die Summenmethode bei Tabulierungen 86

V. **Numerische Integration: Viereckverbesserung.** Seite
§ 64. Aufgabe 89
§ 65—66. Trapezverbesserung 90
§ 67. Rechteckverbesserung 93
§ 68. Vergleichung mit der Summenmethode 94

VI. **Numerische Integration: Mittelwertmethoden.**
§ 69—70. Allgemeine Form der Aufgabe 95
§ 71—72. Gewichte für den Fall einer Potenzreihe 98
§ 73—74. Gewichte von *Cotes*; Regel von *Simpson* 100
§ 75—77. Lösung von *Gauß* 103
§ 78—81. Fall einer trigonometrischen Reihe 107

VII. **Trigonometrische Reihen.**
§ 82—84. Rechnung mit einem regelmäßigen Vieleck 114
§ 85—86. Vereinfachungen und Kontrolle 118
§ 87—88. Zusammensetzung von Vielecken 121
§ 89. Hilfssatz über lineare Systeme 124
§ 90—91. Das unregelmäßige Vieleck 126
§ 92. Schlußkontrollen 129
§ 93—97. Verfahren nach *Tschebyschef* 129
§ 98. Einfluß der Abrundungsfehler 136

VIII. **Rekursionsformeln.**
§ 99. Dreigliedrige Formen 139
§ 100. Rechnung nach vorwärts 140
§ 101—103. Rechnung nach rückwärts 143
§ 104. Ausnahmefälle 147

IX. **Interpolation im weiteren Sinne.**
§ 105—106. Aufgabe 149
§ 107. Methode der kleinsten Quadrate 152
§ 108—110. Verfahren von *Cauchy* 155

Berichtigungen.

S. 55, Z. 11 v. u.: § 43 zu streichen.
S. 85, Z. 4 v. o.: statt § 26 zu lesen § 62.
S. 87, Z. 3 v. o.: statt achtstellige zu lesen zwölfstellige.

Einleitung.

§ 1. Das numerische Rechnen wird in alter Zeit als eine Kunst, nämlich als die Rechenkunst bezeichnet. Gegenwärtig ist es jedoch angemessener, nicht von einer Kunst, sondern von einer Technik des Rechnens zu reden, denn der Sprachgebrauch pflegt heutzutage bei der Ausübung einer Kunst zu unterscheiden zwischen dem geistigen Inhalt als der eigentlichen künstlerischen Leistung und zwischen den Hilfsmitteln, die zur sinnfälligen Darstellung jenes Inhalts benutzt werden. Demgemäß wollen wir sagen, das numerische Rechnen sei eine Technik, deren Zweck in der ziffermäßigen Verfolgung mathematischer Größenbeziehungen besteht. Der Zusatz „ziffermäßig“ ist wesentlich, denn neben dem numerischen Rechnen gibt es auch ein graphisches, das namentlich in den Ingenieurwissenschaften ausgiebig Verwendung findet.

Die Entwicklung der Rechentechnik kann auf eine lange Geschichte zurückblicken, denn die Zahlwörter und die ihnen entsprechenden Zahlbegriffe samt den ersten Ansätzen zu den arithmetischen Grundoperationen sind zweifellos ebenso alt wie die menschliche Sprache. Selbstverständlich ist die Entwicklung bis zu dem heutigen Zustande nur allmählich erfolgt, zum Teil im Anschlusse an die Ausbildung der Arithmetik und Analysis. Ich beschränke mich darauf, die Hauptmomente zu erwähnen; eine eingehende Belehrung findet man in dem umfangreichen Werke von *Moritz Cantor*, *„Vorlesungen über Geschichte der Mathematik“*.

§ 2. Als der erste Schritt in der Entwicklung der Rechentechnik ist zu bezeichnen die sinnfällige Darstellung der Zahlgrößen, sei es durch einen Abakus, sei es durch einfache oder zusammengesetzte Zahlzeichen, wobei es zunächst gleichgültig war, ob diese Zeichen dem Vorrat der gebräuchlichen Buchstaben entlehnt wurden oder aber aus besondern „Ziffern“ bestanden. Wie bekannt haben sich nach und nach die arabischen oder — richtiger gesprochen — indischen Ziffern durchgesetzt; die römische Ziffer dient, soweit sie nicht als archaistischer Schmuck auftritt, nur noch als Nummerzeichen.

Der zweite wichtige Schritt war die Einführung des Nullzeichens, das eine der größten Wohltaten darstellt, die ein Erfinder jemals der Menschheit erwiesen hat, denn nun erst war es möglich, das Prinzip des Stellenwertes konsequent durchzuführen und damit eine Unsumme rechnerischer Arbeitsleistung zu ersparen. Die folgerichtige Weiterentwicklung des Stellenwertprinzips führte dann zu dem Rechnen mit Decimalbrüchen, deren Erfindung an die Namen von *Stevin* und *Joost Bürgi* geknüpft ist, während eine kleine aber nicht unwichtige Äußerlichkeit, nämlich das Decimalkomma oder — nach einer heute weit verbreiteten Druckgewohnheit — der Decimalpunkt von *Kepler* herrührt.

Stevin hat seinerzeit, in klarer Erkenntnis der Wichtigkeit des Gegenstandes, die Forderung erhoben, daß entsprechend der streng decimalen Anlage der Rechnung fortan auch die Maßgrößen decimal abzustufen seien, wenn anders die neuerfundene Rechnungsart ihren vollen Wert entfalten solle. Leider ist die Rechentechnik auch heute noch recht weit davon entfernt, den aus einer Erfüllung der genannten Forderung fließenden Gewinn einheimsen zu können. Denn selbst auf rein wissenschaftlichem Gebiete, wo man noch am ehesten das Durchdringen einer solchen Neuerung erwarten sollte, behauptet sich z. B. in der Winkelteilung ein atavistischer Rest der alten Sexagesimalteilung, die doch nur in einer früheren, unvollkommeneren Entwicklungsphase der Rechentechnik einen vernünftigen Sinn besaß. Zwar sind die Tertien, Quarten, Quinten u. s. w., die sich noch bei *Euler* vorfinden, aus der Literatur verschwunden, desto zäher behaupten sich aber noch die Sekunde und Minute. Dies ist um so mehr zu bedauern, als solche Rudimente früherer Bildungen ein ernsthaftes Hindernis sind, den massenhaften Ziffernverbrauch, der schon jetzt mit seiner routinemäßigen Öde in drückender Weise auf einzelnen rechnenden Wissenschaften lastet, auf besondere Rechenmaschinen abzuwälzen, die eigens für gewisse massenhaft auftretende Operationen konstruiert sind.

Durch die Aufstellung der Decimalbrüche war die Aufgabe gelöst worden, der Gesamtheit der Zahlengrößen eine durch bloße Aneinanderreihung von nur zehn Zeichen gebildete Gesamtheit von Zahlensymbolen Glied für Glied in der Weise zuzuordnen, daß das Operieren mit den Größen durch ein Rechnen, d. h. durch ein Operieren mit den Symbolen ersetzt werden kann. Hierbei kam offenbar derselbe Erfindungsgedanke zur Geltung, der bei der schriftlichen Wiedergabe des gesprochenen Wortes zur Einführung der Lautschrift an Stelle der Silben- und Wortschrift geleitet hat. Mathematisch gesprochen beruhen die Decimalbrüche darauf, daß jede Zahlengröße durch eine endliche oder unendliche, nach fallenden Potenzen von Zehn fortschreitende Reihe dargestellt wird, deren Koeffizienten die in die

einzelnen „Stellen“ des Zahlensymbols einzusetzenden Ziffern angeben. Diese Bemerkung läßt zugleich erkennen, daß die Decimalbrüche nur eine von den unendlich vielen möglichen Lösungen der genannten Aufgabe enthalten. Die gewählte Lösung besitzt jedoch zwei wesentliche Eigenschaften, nämlich Einfachheit und unmittelbare Anpassung an die decimal zählende Sprache, und wird eben dadurch für den regelmäßigen Gebrauch geeignet.

§ 3. Neben der besprochenen Entwicklung, die mit der Erfindung der Decimalbrüche abschließt, setzt schon sehr früh und unabhängig davon die Einführung eines zweiten Hilfsmittels der Rechentechnik ein, nämlich die Darstellung des numerischen Verlaufes einer Funktion durch Tafeln oder, kürzer ausgedrückt, die „Tabulierung“. Weit vor den Logarithmentafeln, deren erstes Erscheinen knapp dreihundert Jahre zurückliegt, stoßen wir auf die trigonometrischen Tafeln, die dazu dienten, aus den astronomischen Winkelmessungen die Koordinaten der Gestirne herzuleiten; die Sehnentafel des *Almagest* ist uns erhalten, und über die noch ältere Sehnentafel *Hipparchs* besitzen wir wenigstens das Zeugnis des Mathematikers *Theon von Alexandria*. Wiederum älter aber sind die Ephemeriden, die nach Ausweis der Keilschriftenfunde von den babylonischen Astronomen angelegt wurden, um auf Grund der vorliegenden Beobachtungsreihen in einer für die damalige Zeit durchaus rationellen Weise die Örter von Sonne und Mond vorauszuberechnen.

Die numerischen Tafeln sind jedoch nicht die einzige Bereicherung, die den Hilfsmitteln der Rechentechnik durch die fortschreitende Entwicklung der Mathematik zu teil geworden ist. Namentlich hat die Ausbildung der Analysis des Unendlichen den Anstoß zur Erfindung von besonderen Methoden gegeben, deren Darstellung den Hauptinhalt der nachfolgenden Abschnitte bilden wird. Man kann sagen, daß es sich hierbei um ein Neuland handelt, dessen Erforschung erst begonnen hat.

§ 4. Will man den gegenwärtigen Inhalt der Rechentechnik mit kurzen Zügen umschreiben, so kann man drei Abschnitte unterscheiden. Der erste Abschnitt umfaßt nach alter Ausdrucksweise die vier Species, deren Handhabung als ein allgemeiner Kulturbesitz anzusehen ist, denn sie findet sich auch da, wo die Kenntnis des Lesens und Schreibens fehlt, nur daß dann gewöhnlich die geschriebene Ziffer durch die Gleitknöpfe eines Abakus ersetzt wird.

Der zweite Abschnitt deckt sich im wesentlichen mit dem Lehrpensum der höheren deutschen Mittelschulen, d. h. er reicht bis zu der Auflösung von Gleichungen und dem Gebrauch der logarithmischen und trigonometrischen Tafeln. Der damit abgegrenzte Bestand

von rechnerischen Hilfsmitteln und Methoden genügt erfahrungsgemäß für den größeren Teil der Aufgaben, denen man bei den Anwendungen der Mathematik begegnet. Wir werden den Inhalt dieses Abschnittes hier in der Hauptsache als bekannt voraussetzen, zumal da der Gegenstand in dem Buche von *Lüroth*, *„Vorlesungen über numerisches Rechnen“* eine angemessene Darstellung gefunden hat.

Der dritte Abschnitt, dessen Inhalt uns im folgenden beschäftigen wird, umfaßt in der Hauptsache Methoden, die ursprünglich aus dem Bedürfnis der rechnenden Astronomie heraus geschaffen worden sind. Von ihnen gilt auch heute noch die *Gauß*sche Bemerkung: „methodi, quae a geometris rarius quam par est in usum vocari videntur“. Es ist z. B. durchaus nicht allgemein bekannt, daß die Auswertung eines bestimmten Integrals sehr oft durch eine mechanische Quadratur rascher erledigt werden kann, als durch eine analytische Entwicklung; noch weniger bekannt ist, daß jene Methoden häufig auch da noch ihre Schuldigkeit tun, wo die gangbaren analytischen Hilfsmittel versagen.

Für die hierher gehörigen Entwicklungen ist von anderer Seite der Name *wissenschaftliches Rechnen* eingeführt worden; ich will denselben mangels einer geeigneteren Bezeichnung einfach beibehalten.

§ 5. Da das Rechnen Zeit und Aufmerksamkeit kostet und deshalb gemeiniglich mit irgend einem Äquivalent bezahlt werden muß, so ist man von jeher bestrebt gewesen, Einrichtungen zu ersinnen, die einerseits Zeit gewinnen lassen, andererseits die durch Nachlassen der Aufmerksamkeit erzeugte Gefahr von Rechenfehlern vermindern. Von solchen Einrichtungen sind hier zu nennen: 1) die graphisch-geometrischen Hilfsmittel, zu denen z. B. der viel gebrauchte logarithmische Rechenschieber und die zum Teil äußerst scharfsinnig erdachten Integrierapparate gehören; 2) die aus Trieb- und Schaltwerken zusammengesetzten Rechenmaschinen im engeren Sinne des Wortes; 3) die numerischen Tafeln. Die erste Gruppe werden wir beiseite lassen, weil wir uns von vornherein auf die rein numerischen Methoden beschränken wollen. Ebenso sollen auch die Rechenmaschinen nicht besonders besprochen werden, obgleich sie recht eigentlich für das Ziffernrechnen bestimmt sind. Einerseits nämlich sind die gegenwärtig verbreiteten Maschinen nur auf die einfachen Operationen

$$a + b, \quad a - b, \quad ab, \quad a + bc, \quad a - bc,$$

aber nicht auf die selbsttätige Berechnung zusammengesetzterer Verbindungen eingerichtet, andererseits läßt sich eine klare Einsicht in die Konstruktion dieser Triebwerke, wenn man die Maschine nicht vor sich hat, nur an der Hand von Zeichnungen durch eine meist nicht sehr kurze Beschreibung vermitteln, während es, wenn man

dem Werk gegenüber sitzt, erfahrungsgemäß nicht schwer ist, sich in die Handhabung hineinzufinden, denn die vorhandenen Konstruktionen sind so eingerichtet, daß ihre Bedienung zwar nicht gedankenlos, wohl aber mit möglichst wenig Nachdenken erfolgen kann. Aus diesem Grunde mögen hier nur einige Bemerkungen mehr allgemeiner Art Platz finden.

Der Gebrauch der Maschine begegnete bei den wissenschaftlichen Rechnern meistens einer mehr oder minder ausgesprochenen Zurückhaltung, die nur zum Teil berechtigt ist. Es steht ja außer Frage, daß der weitaus größere Teil des Ziffernverbrauchs, den gegenwärtig z. B. der Astronom zu bewältigen hat, für die Maschinenarbeit nicht in Betracht kommt: wenn die auftretenden Größen nur zwei oder drei oder höchstens vier geltende Ziffern enthalten, so fördert die herkömmliche „Handarbeit" in Verbindung mit geeigneten Hilfstafeln rascher als die Maschine, zumal der geschulte Rechner dabei in der Lage ist das auszunutzen, was *Gauß* als die „lokalen Vorteile" bezeichnet. Es wäre jedoch falsch, diesen Satz als allgemein gültig anzusehen. So sind, um ein Beispiel anzuführen, in der Abhandlung von *W. G. Höckner*, „Einschaltung von Punkten in ein durch Koordinaten gegebenes trigonometrisches Netz mit ausgiebiger Verwendung einer Rechenmaschine. (Inaugural-Dissertation, Leipzig 1891)" die Vorschriften für die Punkteinschaltung unter der Voraussetzung bearbeitet worden, daß die Rechnung nicht mit den üblichen logarithmischen Tafeln, sondern mit der Maschine unter Benutzung von Tafeln der natürlichen trigonometrischen Funktionen bis zu sieben Decimalen bewerkstelligt werden solle; das Verfahren ist im Jahre 1891 bei der Leipziger Stadtvermessung an Stelle der bis dahin gebrauchten logarithmischen Rechnung eingeführt und seitdem auf Grund der dabei gemachten Erfahrungen beibehalten worden.

Ein anderer Fall, der einfacherer Art ist und sich auf die „verkettete" Multiplikation bezieht, möge hier erwähnt werden, weil später darauf zurückzukommen sein wird. Die *Thomas*-Maschine und ihre Verwandten, deren Einrichtung und Verwendung hier als bekannt vorausgesetzt werden soll, gestattet den Wert eines Aggregats von der Form

$$aA + bB + \cdot\cdot + mM$$

gewissermaßen „vom Blatt" herunterzuspielen, sobald man die Faktoren $a, A, \cdot\cdot$ vor sich hat. Man rechnet zunächst das Produkt aA und läßt es in der Maschine stehen. Dann rechnet man das nächste Produkt bB, welches sofort bei seinem Entstehen von der Maschine zu dem Produkt aA hinzugefügt wird, so daß das Aggregat $aA + bB$ zum Vorschein kommt. In dieser Weise geht es weiter, bis das vorgelegte Aggregat fertig gerechnet ist. Die hier beschriebene Leistung der Maschine läßt sich nun vorteilhaft verwerten, wenn es sich darum

handelt, die Glieder der mit dem gemeinsamen Multiplikandus a gebildeten Produktreihe

$$aA, \quad aB, \cdot\cdot\, aL, \quad aM$$

einzeln zu finden. Statt nämlich die Produkte einzeln und unabhängig voneinander zu bilden, fügt man zunächst der Reihe der Multiplikatoren das Glied 1 (oder auch eine passende Potenz von 10) hinzu, setzt die Größen

$$A, \quad B - A, \quad C - B, \cdot\cdot\, M - L, \quad 1 - M$$

an und berechnet wie vorhin angegeben das Aggregat

$$aA + a(B - A) + \cdots + a(M - L) + a(1 - M),$$

wobei jedes der gesuchten Produkte und als Schlußwert die Zahl a erscheint. Damit ist eine durchgreifende Kontrolle gegeben, welche die zur Bildung der Differenzen $B - A, \cdot\cdot$ nötige Mühe reichlich aufwiegt.

Die Faktoren a, A, B, $\cdot\cdot$ sind vorstehend zunächst als absolute Zahlen zu denken. Das Verfahren bleibt jedoch noch anwendbar, wenn die einzelnen Faktoren mit einem Vorzeichen versehen sind, nur daß dann in der Maschine statt der negativen Produkte dekadische Ergänzungen zum Vorschein kommen, was unter Umständen geradezu vorteilhaft sein kann.

§ 6. Das vorhin betrachtete allgemeine Aggregat

$$aA + bB + \cdot\cdot + mM$$

gibt noch zu einer anderen Bemerkung Anlaß. Die oben beschriebene Rechnung mit der Maschine erscheint auf den ersten Anblick als äußerst vorteilhaft, denn man hat nur die schlechthin notwendigen arithmetischen Operationen auszuführen und braucht keine einzige Zwischengröße niederzuschreiben. In Wirklichkeit liegt jedoch die Sache etwas anders. Soll nämlich die Rechnung, wie es doch nötig ist, kontrolliert werden, so hat man sie im allgemeinen zu wiederholen und, falls die Kontrolle nicht stimmt, eine nochmalige Wiederholung vorzunehmen, da man außer stande ist, den eigentlichen Fehler nachträglich aufzufinden und durch eine Korrektur des Resultats unschädlich zu machen. Aus diesem Grunde wird man es, namentlich wenn ein vielgliedriges Aggregat vorliegt, der Regel nach vorziehen, die Produkte einzeln zu berechnen und einzeln zu kontrollieren. Diese Erschwerung der Rechnung würde jedoch sofort verschwinden, wenn die Maschine mit einer Einrichtung versehen wäre, welche bei der zuerst beschriebenen Rechnungsweise die benutzten Faktoren und die schrittweise entstehenden Produktsummen selbsttätig auf einen Papierstreifen druckt, denn dann hätte man behufs der Kontrolle nur

nachzusehen, ob die richtigen Faktoren in die Maschine gegeben worden sind, und man könnte einen etwa begangenen Fehler nachträglich ohne Neurechnung fortkorrigieren.

Eine „Aggregatmaschine“ der gedachten Art würde voraussichtlich nur für den einen Zweck, nämlich Berechnung von Produktsummen, tauglich sein. Dies wäre jedoch durchaus nicht als ein Mangel anzusehen, denn Produktsummen kommen so überaus häufig vor, daß der Besitz einer brauchbaren Aggregatmaschine für größere Institute zweifellos eine merkliche Arbeitsersparnis bedeuten würde. Die gleiche Bemerkung gilt für gewisse andere Aufgaben. Einige von den Wissenschaften, die mit einem starken Ziffernverbrauch belastet sind, wie z. B. die Fixstern-Astronomie, stehen — falls sie ihre Probleme mit Erfolg bewältigen wollen — vor der Notwendigkeit, über kurz oder lang zu einem zentralisierten Großbetrieb überzugehen, bei dem bestimmte zusammengesetzte Rechenoperationen hunderttausend und millionenfach auszuführen sind. Daß hierbei die Maschine, und zwar die jedesmal für eine einzige bestimmte Aufgabe möglichst vollkommen durchgebildete Maschine, der unentbehrliche Helfer sein wird, steht meiner Überzeugung nach außer Frage. Es besitzt übrigens ein gewisses psychologisches Interesse, zu beobachten, wie auf der einen Seite z. B. bei der Ausrüstung einer Sternwarte für ein neues Instrument ansehnliche Summen verausgabt und dabei nach Möglichkeit die Fortschritte der Technik herangezogen werden, wie man aber auf der andern Seite vorläufig gar nicht daran denkt, ob die Technik, die so sinnreiche Apparate wie die Schreibmaschine und die Kontrollkasse geschaffen hat, nicht etwa auch für die Rechenarbeit, welche zeitlich bemessen ein Vielfaches der Beobachtungsarbeit ausmacht, hilfreiche Dienste leisten könnte.

§ 7. Unter den technischen Erleichterungen war oben an dritter Stelle die Tabulierung genannt worden. Die Wichtigkeit dieses Hilfsmittels erhellt am deutlichsten aus dem Umfange der darauf bezüglichen Literatur. Würde man die gedruckten Tafelausgaben an einem Orte vereinigen, so entstünde eine ganz ansehnliche Bibliothek, in der die drei großen Ephemeridenreihen des „Berliner Jahrbuchs“, der „Connaissance des temps“ und des „Nautical Almanac“ für sich allein zusammengenommen fast fünfhundert Bände zählen.

Da eine Tafel den Zweck verfolgt, das Rechnen zu erleichtern, so ist sie selbstverständlich derart einzurichten, daß der gewollte Zweck möglichst vollständig erreicht wird, daß also im besondern die gesuchten Zahlen möglichst rasch gefunden werden können. Daraus fließen für das Äußere der Tafeln gewisse Forderungen, die hier kurz berührt werden mögen, da es — wie die Erfahrung lehrt — gut ist, immer wieder daran zu erinnern.

Besitzt eine für den täglichen Gebrauch bestimmte Tafel übergroßes Format, so wird sie unhandlich. Wie lästig das wird, kommt dem Rechner recht deutlich zum Bewußtsein, wenn er genötigt ist, einen Folianten, wie z. B. den großen zehnstelligen Thesaurus logarithmorum von *Vega*, zu wälzen. Man kann sagen, daß das sogenannte Lexikonoktav oder, wenn es sich nicht vermeiden läßt, ein gleich hohes Quart etwa die Grenze bezeichnet, von der ab die Unhandlichkeit beginnt. Auf der andern Seite wird aber auch ein sehr kleines Format unhandlich und zwingt — was noch nachteiliger ist — zur Wahl einer sehr kleinen Ziffer, beeinträchtigt also die rasche Lesbarkeit der Zahlen. Vergleicht man Tafeln, die verschieden großen Druck aufweisen, hinsichtlich der raschen Lesbarkeit miteinander, so ergibt sich, daß die sogenannte Petitziffer die Grenze bezeichnet, unter die man nur ausnahmsweise heruntergehen darf; der ausgedehnte Gebrauch einer noch kleineren Ziffer wirkt auf einen Leser mit normaler Sehweite geradezu als Augenpulver.

Neben der Größe der Ziffer ist des weitern ihre Form von Wichtigkeit. Bis zum Anfange des achtzehnten Jahrhunderts war so gut wie ausschließlich eine Form im Gebrauch, die gegenwärtig als mediävale Ziffer bezeichnet wird. Charakteristisch für dieselbe sind die nach oben und unten ungleiche Höhe der zehn Zeichen, die rundliche Gestalt und die verhältnismäßig schwache Schattierung. Später nahm ein veränderter Geschmack an der ungleichen Höhe Anstoß und führte statt dessen Ziffern von durchweg gleicher Höhe, schlankerer Gestalt und starker Schattierung ein. Das Eindringen dieser Mode in die Tafelwerke bedeutete eine ausgesprochene Verschlechterung, denn die erwähnten Merkmale der mediävalen Ziffer begünstigen das rasche Lesen der Zahlen.

Die Menge der auf einer Tafelseite unterzubringenden Ziffern macht es notwendig, den bedeckten Raum durch geeignete Mittel derart zu gliedern, daß das Auge rasch zu der Stelle hingleitet, wo die verlangte Zahl zu suchen ist. Für die Gliederung der Spalten sind im allgemeinen vertikale Linien zu bevorzugen, und zwar der Raumersparnis halber, denn eine Gliederung durch bloßes Auseinanderrücken der Spalten wirkt äußerst unruhig, sobald die trennenden freien Räume nicht sehr reichlich bemessen sind. Etwas anders liegt die Sache bei der Gliederung der Zeilen. Hier empfiehlt es sich in der Regel, die größeren Abteilungen durch Linien, die kleineren durch vermehrten Zeilenabstand hervorzuheben. Versucht man die Gliederung nur durch verstärkten „Durchschuß“ zu bewirken, so bedarf es, wie vorhandene Beispiele lehren, besonderer Aufmerksamkeit, um das quer über die Seite wandernde Auge ohne Zuhilfenahme des Fingers auf einer Zeile festzuhalten.

Es versteht sich von selbst, daß neben den genannten Stücken

bei einer guten Tafelausführung auch noch andere Dinge zu beachten sind, wie z. B. die angemessene Hervorhebung gewisser Ziffern, die Überschriften der Spalten und Seiten, die Unterbringung der etwa nötigen Nebentäfelchen, die Beschaffenheit des Papiers, die Dunkelheit des manchmal übermäßig hell gehaltenen Drucks und dergleichen mehr. Für diese Dinge lassen sich jedoch nicht gut bestimmte Regeln aufstellen; man kann allgemein nur sagen, daß als oberste Richtschnur immer die tunlichste Schonung der Augen und der Aufmerksamkeit zu dienen hat, denn ein Verstoß hiergegen bedeutet jederzeit zugleich einen Angriff auf die Augen und die Nerven des Nebenmenschen.

§ 8. Die nachstehende Darstellung der Methoden des wissenschaftlichen Rechnens wird sich folgendermaßen gliedern. Den Anfang bilden gewisse Eigenschaften der Differenzen und Summen, sowie — im Zusammenhange damit — der numerischen Tafeln. Die Nutzanwendung dieser Eigenschaften führt zu der Interpolation im engeren Sinne, ferner zu der numerischen Differentiation und zu gewissen Methoden der mechanischen Quadratur oder, wie man jetzt besser sagt, der numerischen Integration. Daran werden sich die übrigen gebräuchlichen Methoden der numerischen Integration und die Behandlung der trigonometrischen Reihen anfügen. Den Schluß soll endlich je ein Abschnitt über Rekursionsformeln und über Interpolation im weiteren Sinne bilden. Auf die kleinen Kunstgriffe, die in der Rechentechnik nach und nach ausfindig gemacht worden sind, werde ich nur in einzelnen Fällen eingehen, da sich diese Dinge durch mündliche Unterweisung besser übermitteln lassen als durch schriftliche.

Um später den Zusammenhang nicht zu unterbrechen, mögen endlich am Schlusse dieser Einleitung einige Bemerkungen zusammengestellt werden. Die natürlichen und die gemeinen Logarithmen sollen durch die Zeichen log und Log unterschieden werden. Für negative Logarithmen wird nach einer zweckmäßigen, aber selbst heute noch nicht allgemein durchgedrungenen Regel ihre dekadische Ergänzung geschrieben, also z. B. 8.736 — 10 statt — 1.264, wobei der Substrahendus — 10 fortgelassen wird, sobald kein Mißverständnis zu befürchten ist. Ist der Numerus negativ, so wird dieser Umstand nach dem Vorgange von *Gauß* durch ein dem Logarithmus angehängtes n angezeigt.

Zur Erleichterung des Drucks wird der Quotient „a dividiert durch b" häufig in der Form $a:b$ geschrieben werden. Aus demselben Grunde werde ich für die Exponentialfunktion e^z das Zeichen $\exp z$ anwenden, das nach dem Muster der Zeichen sin, cos u. s. w. gebildet ist.

Bedeuten die Größen $a_1, a_2, \cdots$ feste Koeffizienten, und ist zur Abkürzung

$$A = a_1 + a_2 + \cdots,$$

so soll der mit den Veränderlichen $x_1, x_2, \cdots$ gebildete Quotient

$$M = (a_1 x_1 + a_2 x_2 + \cdots) : A$$

als das *mit den Gewichten a gebildete Mittel der x* bezeichnet werden. Der Ausdruck stammt ursprünglich aus der Statik und ist dann später von der Ausgleichungsrechnung aufgenommen worden. Besitzen die a einerlei Vorzeichen, so gehört M sicher der Wertstrecke an, die von dem größten und kleinsten unter den x begrenzt wird, dagegen kann M außerhalb dieser Strecke liegen, wenn die a ungleiches Vorzeichen haben. Der Bequemlichkeit halber werden wir auch für den letzteren Fall den Ausdruck *Mittel* zulassen.

In dem Aggregat $a_1 x_1 + a_2 x_2 + \cdots$ sind, wie wir öfter sagen werden, die Größen *x durch die Multiplikatoren a miteinander verbunden.* Der Ausdruck läßt sich offenbar auch auf Gleichungen anwenden, indem man unter den x erst die linken, dann die rechten Seiten der Gleichungen versteht.

Erster Abschnitt.

Differenzen und Summen.

§ 9. Den ersten Gegenstand der Untersuchung sollen wie angegeben gewisse Eigenschaften der Differenzen- und Summenreihen bilden. Hierbei knüpfen wir an die nachstehende Tabelle an, deren Einrichtung zunächst zu erläutern ist.

Tabelle I.

$$\begin{array}{cccc}
a_1 & & & \\
 & b_1 & & \\
a_2 & & c_1 & \\
 & b_2 & & d_1 \\
a_3 & & c_2 & \\
 & b_3 & & d_2 \\
a_4 & & c_3 & \\
\cdot\;\cdot\;\cdot & \cdot\;\cdot\;\cdot & \cdot\;\cdot\;\cdot & \cdot\;\cdot\;\cdot
\end{array}$$

Die in der ersten Spalte enthaltenen Größen a sollen Zahlen bedeuten, die irgendwie gegeben sind. Aus ihnen entstehen die Größen b, indem man jedes a von dem folgenden abzieht und die Differenz in die zweite Spalte auf die Mitte zwischen Subtrahendus und Minuendus setzt. Die derart gefundenen Größen b werden dann in genau derselben Weise behandelt wie die Größen a und liefern dadurch die Größen c der dritten Spalte, und so geht es weiter zu den folgenden Spalten hin. Hiernach ist der Zusammenhang zwischen den einzelnen Gliedern der Tabelle durch das Gleichungssystem

$$a_2 - a_1 = b_1, \quad a_3 - a_2 = b_2, \quad a_4 - a_3 = b_3, \cdots \qquad (1)$$

$$b_2 - b_1 = c_1, \quad b_3 - b_2 = c_2, \quad b_4 - b_3 = c_3, \cdots \qquad (2)$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

gegeben. Die Größen $b, c, d, \cdots$ tragen den Namen *Differenzen erster, zweiter, dritter u. s. w. Ordnung* oder kürzer *erste, zweite, dritte u. s. w. Differenzen* der Größen a. Die ganze Tabelle bildet sonach ein Differenzensystem oder wie es fortan heißen soll ein *Differenzenschema*, wobei man der Gleichförmigkeit halber die Ausgangsgrößen a als Differenzen nullter Ordnung bezeichnen kann.

Ist die Anzahl der a begrenzt, z. B. gleich n, so enthält jede Spalte ein Glied weniger als die unmittelbar vorhergehende. Da ferner

die Summe aus Gliederzahl und Ordnungsnummer für alle Spalten denselben Wert n annimmt, so enthält die Spalte mit der Ordnungsnummer $n - 1$ nur ein Glied: das Schema bricht also mit dieser Spalte ab.

Summiert man die Gleichungen (1), so erkennt man, daß die Summe der Größen b gleich der Differenz der beiden äußersten a ist. Ebenso folgt aus den Gleichungen (2), daß die Summe der c gleich der Differenz der beiden äußersten b ist, und so fort. Diese Bemerkung liefert eine bequeme und nützliche Kontrolle für die Berechnung der Differenzen.

§ 10. Die Tabelle I läßt sich nach links hin durch Vorlegen neuer Spalten derart erweitern, daß die erweiterte Tabelle wiederum ein Differenzenschema bildet. Zu dem Ende legen wir der Spalte der a zunächst eine Spalte vor, deren Glieder $u_1, u_2, \cdot\cdot$ den Gleichungen

$$u_2 - u_1 = a_1, \quad u_3 - u_2 = a_2, \quad u_4 - u_3 = a_3, \cdots$$

genügen. Hieraus ergeben sich die Gleichungen

$$\begin{aligned} u_2 - u_1 &= a_1, \\ u_3 - u_1 &= a_1 + a_2, \\ u_4 - u_1 &= a_1 + a_2 + a_3, \\ &\cdots\cdots\cdots \end{aligned}$$

Man erhält also, wenn das Anfangsglied u_1 irgendwie festgesetzt worden ist, die übrigen u durch schrittweises Aufsummieren der a. Aus diesem Grunde werden die u als die zu den a gehörigen *Summengrößen erster Ordnung* bezeichnet. In derselben Weise kann man nach Bildung der u eine neue Spalte vorlegen, deren Glieder $v_1, v_2, \cdot\cdot$ den Bedingungen

$$v_2 - v_1 = u_1, \quad v_3 - v_2 = u_2, \quad v_4 - v_3 = u_3, \cdots$$

genügen und nach Festsetzung des Anfangsgliedes v_1 aus den Gleichungen

$$\begin{aligned} v_2 - v_1 &= u_1, \\ v_3 - v_1 &= u_1 + u_2, \\ v_4 - v_1 &= u_1 + u_2 + u_3, \\ &\cdots\cdots\cdots \end{aligned}$$

durch Aufsummieren der u erhalten werden. Die v werden dann als die zu den a gehörigen *Summengrößen zweiter Ordnung* bezeichnet. Dieser Prozeß des Anfügens nach links hin läßt sich offenbar beliebig weit fortsetzen, weil ein Abbrechen dadurch ausgeschlossen ist, daß jede neue Spalte ein Glied mehr enthält, als die unmittelbar vorhergehende.

Da die Anfangsglieder der Summenspalten irgendwie festgesetzt werden können, so spielen sie die Rolle von vorläufig willkürlichen

Konstanten. Demnach enthält die Summenspalte n^{ter} Ordnung n solcher willkürlicher Summationskonstanten.

Da die nach links hin erweiterte Tabelle beständig den Charakter eines Differenzenschemas bewahrt, so kann man die Summengrößen als Differenzen auffassen, denen die Ordnungsnummern $-1, -2, -3, \cdots$ zukommen.

§ 11. Die einzelnen Glieder des Differenzenschemas entstehen aus den Größen nullter Ordnung nur durch Addition und Subtraktion. Infolgedessen werden innerhalb des Schemas lineare Beziehungen stattfinden, die jetzt aufgesucht werden sollen.

Die Spalte nullter Ordnung enthalte zunächst nur die n Glieder $a_1, \cdots a_n$. Diese Reihe werde jetzt nach vorwärts und rückwärts verlängert, wobei die neu hinzugefügten a der Einfachheit halber aus lauter Nullen bestehen sollen. Setzt man nun mit der willkürlichen Größe t die Reihe

$$S = \sum_p a_p t^p$$

an, in der p von $-\infty$ bis $+\infty$ läuft, so wird

$$S(1-t) = \sum_p (a_p - a_{p-1})\, t^p = \sum_p b_{p-1} t^p .$$

Die Differenzen erster Ordnung treten also als die Koeffizienten der Reihe auf, die durch Multiplikation von S mit $1-t$ entsteht. Wiederholt man diese Operation immer von neuem, so ergibt sich, daß die Differenzen q^{ter} Ordnung als Koeffizienten der Reihe erhalten werden, welche entsteht, wenn man die Reihe S mit der binomischen Entwicklung

$$(1-t)^q = (q)_0 - (q)_1 t + (q)_2 t^2 - \cdots$$

multipliziert. Demnach besitzen die Differenzen q^{ter} Ordnung die Gestalt

$$(q)_0 a_{p+q} - (q)_1 a_{p+q-1} + (q)_2 a_{p+q-2} - \cdots, \tag{3}$$

wo die Reihe mit dem Gliede $\pm (q)_q a_p$ von selber abbricht.

Der gefundene Satz läßt sich sofort auf die Summengrößen übertragen. Sieht man die ersten Summen als Differenzen der Ordnung -1 an und multipliziert demgemäß die Reihe S mit der Reihe

$$(1-t)^{-1} = 1 + t + t^2 + \cdots,$$

so liefern die Koeffizienten der dadurch entstehenden Reihe in der Tat die oben mit u bezeichneten Summengrößen, sobald man festsetzt, daß das Glied u_1 und folgeweise auch die vorhergehenden u Größen gleich Null sein sollen. Wie sich hieraus sofort weiter ergibt, ent-

stehen dann die Summengrößen r^{ter} Ordnung, wenn man S mit der binomischen Entwicklung

$$(1 - t)^{-r} = (r - 1)_0 + (r)_1\, t + (r + 1)_2\, t^2 + \cdots$$

multipliziert und überdies festsetzt, daß die Summengrößen mit dem Index 1 jedesmal null sein sollen.

§ 12. Wenn in der betrachteten Tabelle die a-Spalte n Glieder $a_1, \cdots a_n$ enthält, so laufen die ersten Differenzen von b_1 bis b_{n-1}, die zweiten Differenzen von c_1 bis c_{n-2} u. s. w. Denkt man sich nun die Tabelle vollständig hingeschrieben und dann eine neue Tabelle angesetzt, deren erste Spalte dieselben Größen a, aber in umgekehrter Reihenfolge enthält, so besitzt das neue Differenzenschema, wie leicht zu erkennen ist, folgendes Aussehen:

a_n			
	$-b_{n-1}$		
a_{n-1}		c_{n-2}	
	$-b_{n-2}$		$-d_{n-3}$
a_{n-2}		c_{n-3}	
	$-b_{n-3}$		$-d_{n-4}$
. . .	. . .	. . .	. . .

Die vorgenommene Änderung läßt sich so auffassen, als ob die erste Spalte der ursprünglichen Tabelle um eine Zeile herum umgeklappt worden sei; man kann also sagen, daß das *Umklappen* in allen Spalten des Schemas die Reihenfolge der Glieder und in den ungeraden Spalten überdies die Vorzeichen umkehrt.

§ 13. Außer der Tabelle I sei jetzt noch ein zweites Differenzenschema angesetzt worden, dessen Glieder die nachstehende Tabelle ersichtlich macht:

Tabelle II.

A_1			
	B_1		
A_2		C_1	
	B_2		D_1
A_3		C_2	
. . .	. . .	. . .	. . .

Beide Tabellen verbinden wir jetzt Glied für Glied vermittels der beiden festen Multiplikatoren g und G, sodaß das Größensystem

$$U_p = g a_p + G A_p, \quad V_p = g b_p + G B_p, \quad \text{u. s. w.}$$

entsteht. Werden nun die Größen $U, V, \cdots$ in derselben Weise wie die $a, b,$ und $A, B, \cdots$ tabellarisch zusammengestellt, so entsteht

wiederum ein Differenzenschema. Zum Beweise genügt es die allerersten Glieder zu betrachten. Es ist

$$\begin{aligned} U_2 - U_1 &= (g a_2 + G A_2) - (g a_1 + G A_1) \\ &= g(a_2 - a_1) + G(A_2 - A_1) \\ &= g b_1 + G B_1, \end{aligned}$$

woraus $U_2 - U_1 = V_1$ folgt, d. h. V_1 ist in der Tat die zu U_1 und U_2 gehörigen Differenz.

Das gefundene Ergebnis besagt offenbar, *daß die Differenzen einer Verbindung entstehen, wenn die Verbindung an den Differenzen der Komponenten vollzogen wird.* Dieser Satz, der zunächst nur für zwei Komponenten nachgewiesen ist, gilt wie leicht zu sehen auch für Verbindungen aus beliebig vielen Komponenten.

Zur Erläuterung mögen nachstehend zwei einfache Fälle behandelt werden.

§ 14. Die Tabelle II enthalte, wie jetzt angenommen werden soll, im wesentlichen dieselben Zahlen wie die Tabelle I, jedoch um eine Zeile verschoben, so daß — von Anfang und Ende der Tabellen abgesehen — die Beziehungen

$$A_p = a_{p+1}, \quad B_p = b_{p+1}, \quad \text{u. s. w.}$$

bestehen. Weiter sollen die beiden Multiplikatoren g, G den besondern Wert $\frac{1}{2}$ besitzen. Dann nehmen die aus der Verbindung der beiden Tabellen entspringenden Größen die Gestalt

$$U_p = \tfrac{1}{2}(a_p + a_{p+1}), \quad V_p = \tfrac{1}{2}(b_p + b_{p+1}), \quad \text{u. s. w.}$$

an. Die vorstehenden Gleichungen besagen, daß man wiederum ein Differenzenschema erhält, wenn man in Tabelle I aus je zwei untereinander stehenden Größen das arithmetische Mittel bildet. Diese *Zwischengrößen* werden weiterhin von Nutzen sein.

Eine andere Anwendung des Verbindungssatzes ist folgende. Die Zahlen $A, B, \cdot\cdot$ seien wieder unabhängig von den $a, b, ..$ angesetzt und als Multiplikatoren die Werte $g = 1$ und $G = 1$ gewählt worden, so daß es sich jetzt um die Verbindungen

$$U_p = a_p + A_p, \quad V_p = b_p + B_p, \quad \text{u. s. w.}$$

handelt. Man denke sich nun, daß eine Rechnung mit den Größen a durchgeführt werden solle, daß man aber diese Größen nicht genau kenne, sondern statt dessen nur die mit den Fehlern A behafteten Zahlen $U = a + A$. Dann gilt, wenn man die Differenzen der Größen U, a, A betrachtet, sofort der Satz: *Die Fehler der Differenzen sind zugleich die Differenzen der Fehler.*

An der Hand der vorstehenden Bemerkung läßt sich die Wirkung eines einzelnen Fehlers bequem verfolgen. Wir nehmen an, daß in der ersten Spalte des Schemas nur ein Glied den Fehler x besitze, daß dagegen die übrigen Glieder der Spalte fehlerfrei angesetzt worden seien. Dann enthält die Tabelle der Fehlergrößen in der ersten Spalte das eine von Null verschiedene Glied x, sonst aber nur Nullen. Demnach nimmt das Differenzenschema der Fehler die nachstehende Gestalt an:

Tabelle III.

. . .	. . .	. . .	. . .	. . .
0		0		x
	0		x	
0		x		$-4x$
	x		$-3x$	
x		$-2x$		$6x$
	$-x$		$3x$	
0		x		$-4x$
	0		$-x$	
0		0		x
. . .	. . .	. . .	. . .	. . .

Die Faktoren, mit denen x innerhalb einer Spalte multipliziert wird, sind nichts anderes als die gewöhnlichen Binomialkoeffizienten mit alternierenden Vorzeichen. Der Beweis für diese Bemerkung ergibt sich aus der Formel § 11 (3), wenn man dort bei festgehaltenem q der Zahl p nach und nach alle zulässigen Werte erteilt und dabei beachtet, daß im vorliegenden Falle alle a bis auf eines gleich Null sind.

Aus dem Verhalten der Zahlen in Tabelle III ist zu entnehmen, daß die Wirkung des Fehlers x an Ausbreitung und Größe zunimmt, je weiter man mit der Bildung der Differenzenreihen fortschreitet, und daß ferner die stärkste Fehlerwirkung in jeder Spalte stets in oder an derjenigen Zeile stattfindet, auf welcher der Fehler in der ersten Spalte begangen worden ist.

§ 15. Bevor wir mit der Untersuchung des Differenzenschemas fortfahren, wird es zweckmäßig sein, zum besseren Verständnis der weiterhin behandelten Fragen gewisse Bemerkungen über die numerischen Tafeln vorauszuschicken.

Der normale Weg, eine Funktion $f(x)$ zu tabulieren, setzt sich aus folgenden Schritten zusammen. Zunächst wird für die unabhängige Veränderliche oder das sogenannte *Argument* x die Wertreihe

$$\cdots, a-2h, \quad a-h, \quad a, \quad a+h, \quad a+2h, \cdots$$

angesetzt, deren Glieder mit den konstanten Zuwachs h, dem sogenannten *Tafelintervall*, fortschreiten. Darauf werden zu der vor-

stehenden Argumentreihe die Funktionswerte berechnet. Endlich stellt man die zusammengehörigen Werte von Argument und Funktion tabellarisch zusammen.

Es kommt zuweilen vor, daß innerhalb einer Tafel das Intervall h geändert wird; beispielsweise pflegt man in den Tafeln für die Logarithmen von $\sin x$ und $\operatorname{tg} x$ bei den kleinen Winkeln ein engeres Intervall zu benutzen, als in den übrigen Teilen der Tafel. Solche Fälle können jedoch als eine Vereinigung von mehreren Tafeln aufgefaßt werden, deren jede in sich nur konstantes Intervall besitzt. Aus diesem Grunde darf man sich darauf beschränken, das Tafelintervall als durchweg konstant vorauszusetzen.

Als Beispiel diene das folgende Täfelchen, das aus einer siebenstelligen Tafel der gemeinen Logarithmen ausgeschrieben ist.

Tabelle IV.

x	Log x	I	II	III	IV
50	1.6989700				
		+ 86002			
51	1.7075702		— 1671		
		+ 84331		+ 66	
52	1.7160033		— 1605		— 8
		+ 82726		+ 58	
53	1.7242759		— 1547		— 1
		+ 81179		+ 57	
54	1.7323938		— 1490		— 3
		+ 79689		+ 54	
55	1.7403627		— 1436		
		+ 78253			
56	1.7481880				

Unter x sind die Argumentwerte, unter Log die zugehörigen Funktionswerte, ferner unter den Überschriften I, II, III, IV die Differenzen bis zur vierten Ordnung zusammengestellt. Die Differenzen sind, um unnötigen Nullenballast zu vermeiden, nur mit den geltenden Ziffern oder, anders ausgedrückt, in Einheiten der siebenten Decimale angesetzt.

Betrachtet man das obige Täfelchen näher, so fällt sofort das Verhalten der Differenzen auf: die numerischen Beträge zeigen von links nach rechts eine ausgesprochene Abnahme. Infolgedessen besitzen die Zahlen in senkrechter Richtung einen ausgesprochenen *Gang*, der noch deutlicher zum Ausdruck kommt, wenn man den Inhalt der Tabelle graphisch darstellt. Da ein solches Verhalten bei den numerischen Tafeln durchaus gewöhnlich ist, so entsteht die Frage nach dem allgemeinen Grunde dieser Erscheinung.

§ 16. Bei der Behandlung der gestellten Frage soll für das Differenzenschema der numerischen Tafeln eine Bezeichnungsweise benutzt werden,

die ihrem Wesen nach auf *Gauß* zurückgeht, den Rechnern jedoch zuerst durch die Mitteilungen geläufig wurde, die *Encke* in verschiedenen Bänden des Berliner Jahrbuchs darüber gemacht hat[1]). Das Schema nimmt damit eine Gestalt an, wie sie die nachstehende, sogleich näher zu erläuternde Tabelle zeigt, die als ein Ausschnitt anzusehen ist.

Tabelle V.

Arg.	(− 2)	(− 1)	(0)	(1)	(2)	
$a-2h$	$(a-2, -2)$		$(a-2, 0)$		$(a-2, 2)$	
		$(a-\frac{3}{2}, -1)$		$(a-\frac{3}{2}, 1)$		...
$a-h$	$(a-1, -2)$		$(a-1, 0)$		$(a-1, 2)$	
		$(a-\frac{1}{2}, -1)$		$(a-\frac{1}{2}, 1)$		...
a	$(a, -2)$		$(a, 0)$		$(a, 2)$	
		$(a+\frac{1}{2}, -1)$		$(a+\frac{1}{2}, 1)$		...
$a+h$	$(a+1, -2)$		$(a+1, 0)$		$(a+1, 2)$	
		$(a+\frac{3}{2}, -1)$		$(a+\frac{3}{2}, 1)$		...
$a+2h$	$(a+2, -2)$		$(a+2, 0)$		$(a+2, 2)$	

Die erste Spalte enthält unter der Überschrift „Arg." die Argumentwerte der Tafel. Die Zeilen auf denen die Argumente stehen, nennen wir *Hauptzeilen*, wogegen die Zeilen zwischen den benachbarten Argumenten *Zwischenzeilen* heißen sollen. In der Spalte (0) ist auf jeder Hauptzeile der Funktionswert eingetragen, der zu dem Argument der betreffenden Zeile gehört. Rechterhand von (0) sind unter den Überschriften (1), (2), · · die Differenzen erster, zweiter u. s. w. Ordnung angesetzt, während linkerhand unter den Überschriften (− 1), (− 2), · · die Summengrößen erster, zweiter u. s. w. Ordnung stehen. Die einzelnen Glieder des Schemas sind durch Symbole von der Form $(a+p,\ q)$ bezeichnet. Von den beiden Bestandteilen $a+p$ und q des Symbols ist der zweite, nämlich die Zahl q, innerhalb einer Spalte konstant und stimmt, wie ein Blick auf die Tabelle lehrt, mit der Ordnungsnummer der angesetzten Differenzen überein. Der andere Bestandteil $a+p$ soll *Zeilennummer* heißen; er ist, wie die Tabelle erkennen läßt, längs einer Zeile konstant. In den geraden Spalten, deren Glieder immer auf Hauptzeilen stehen, wird die Zeilennummer aus dem Argument der Zeile erhalten, wenn man in dem Argument den Buchstaben h durch die Zahl 1 ersetzt. In den ungeraden Spalten dagegen, deren Glieder immer auf Zwischenzeilen stehen, ist die Zeilennummer das arithmetische Mittel aus den Nummern der beiden anliegenden Hauptzeilen. Hiernach ist in $(a+p,\ q)$ die Zahl p ein

1) Die auf das wissenschaftliche Rechnen bezüglichen Aufsätze von *Encke*, die auch heute noch die Beachtung des Rechners verdienen, sind durch die von *H. Gravelius* besorgte Ausgabe *„Gesammelte mathematische und astronomische Abhandlungen von J. F. Encke. Berlin 1888"* für Nichtastronomen bequemer zugänglich geworden.

gerades oder ungerades Vielfaches von $\frac{1}{2}$, je nachdem q gerade oder ungerade ist. Daraus folgt, daß der Ausdruck $2p - q$ stets eine gerade Zahl liefert.

Die Tabelle V bezeichnen wir als die *Tabelle der Hauptwerte* oder auch kurz als die *Haupttabelle*. Daneben soll nun sogleich noch eine zweite Tabelle, nämlich die *Tabelle der Zwischenwerte* oder die *Zwischentabelle* angelegt werden. Trägt man in Tabelle V zwischen je zwei untereinander stehenden Gliedern das arithmetische Mittel dieser beiden Glieder ein, so bilden nach § 14 diese Zwischenwerte für sich wieder ein Differenzenschema, dessen Glieder wie vorhin in unzweideutiger Weise durch Zeilen- und Ordnungsnummer vermittelst des Symbols $(a + p, q)$ bezeichnet werden können, nur daß jetzt die Differenz $2p - q$ stets ungerade ist. Löscht man in dem betrachteten Schema die Hauptwerte aus, so erzeugen die stehenbleibenden Zwischenwerte die nachfolgende ohne weiteres verständliche Tabelle.

Tabelle VI.

Arg.	(—2)	(—1)	(0)	(1)	(2)	
$a-h$		$(a-1, -1)$		$(a-1, 1)$		...
	$(a-\frac{1}{2}, -2)$		$(a-\frac{1}{2}, 0)$		$(a-\frac{1}{2}, 2)$	
a		$(a\quad, -1)$		$(a\quad, 1)$		...
	$(a+\frac{1}{2}, -2)$		$(a+\frac{1}{2}, 0)$		$(a+\frac{1}{2}, 2)$	
$a+h$		$(a+1, -1)$		$(a+1, 1)$		...

Die große Übersichtlichkeit der gewählten Bezeichnungsweise beruht offenbar darauf, daß als Zeilennummern die Vielfachen von $\frac{1}{2}$ gewählt worden sind.

§ 17. Da die Werte der in den beiden Tabellen V und VI auftretenden Symbole $(a + p, q)$ ein Differenzenschema erzeugen, so besteht zwischen diesen Symbolen die Beziehung

$$(a + p, q) - (a + p - 1, q) = (a + p - \tfrac{1}{2}, q + 1). \tag{4}$$

Versucht man nun, an der Hand der vorstehenden Gleichung die analytischen Ausdrücke der Größen $(a + p, q)$ für irgend eine Funktion $f(x)$ wirklich zu bilden, so kommt man im allgemeinen auf verwickelte Formen. Eine Ausnahme hiervon macht die Exponentialfunktion. Bezeichnet man mit A und B zwei von den Nummern p und q unabhängige Größen und führt die Abkürzungen

$$\exp(\tfrac{1}{2}h) - \exp(-\tfrac{1}{2}h) = S, \quad \exp(\tfrac{1}{2}h) + \exp(-\tfrac{1}{2}h) = 2C \tag{5}$$

ein, so wird die Bedingung (4) erfüllt, wenn man unter vorläufiger Trennung der Haupt- und Zwischenwerte ansetzt

Hauptwerte, $2p-q$ gerade:

$$(a+p, q) = A \exp(a+ph) S^q, \tag{6}$$

Zwischenwerte, $2p-q$ ungerade:

$$(a+p, q) = B \exp(a+ph) S^q. \tag{7}$$

Wegen des Zusammenhanges zwischen den beiden Arten von Größen besteht nun für gerade Werte von $2p-q$ die Gleichung

$$(a+p, q) + (a+p+1, q) = 2(a+p+\tfrac{1}{2}, q).$$

Setzt man hierin die aus (6) und (7) folgenden Ausdrücke ein, so erhält man die Gleichung

$$A \exp(a+ph) S^q + A \exp(a+ph+h) S^q = 2B \exp(a+ph+\tfrac{1}{2}h) S^q,$$

aus der $\qquad A + A \exp h = 2B \exp(\tfrac{1}{2}h)$

oder wegen (5) $\qquad 2AC = 2B$

folgt. Setzt man also $A = 1$ und demgemäß $B = C$, so können wir sagen: die Ausdrücke

$$(a+p, q) = \exp(a+ph) S^q, \quad (2p-q \text{ gerade}) \tag{8}$$

liefern eine Haupttabelle, deren Zwischenwerte durch die Ausdrücke

$$(a+p, q) = \exp(a+ph) CS^q, \quad (2p-q \text{ ungerade}) \tag{9}$$

gegeben sind. Die Funktionswerte erscheinen hierbei in der Gestalt

$$f(a+ph) = (a+p, 0) = \exp(a+ph);$$

das betrachtete Schema ist also in der Tat mit der Funktion $f(x) = \exp x$ gebildet worden. Diese Bemerkungen werden uns sogleich von Nutzen sein.

§ 18. Zu der oben am Schlusse von § 15 gestellten Frage zurückkehrend machen wir jetzt die Voraussetzung, daß die tabulierte Funktion $f(x)$ in der Umgebung der Stelle $x = a$ nach dem *Taylor*schen Satze durch die Reihe

$$f(a+th) = f_0 + f_1 th + \tfrac{1}{2} f_2 t^2 h^2 + \cdots \tag{10}$$

dargestellt werden könne, in der die Größen $f_0, f_1, f_2, \cdots$ die successiven Ableitungen der Funktion $f(x)$ für $x = a$ bedeuten. Weiter werde die ganze Zahl T und die Größe H so festgesetzt, daß die Reihe (10) noch konvergiert, wenn th gleich TH gemacht wird. Endlich werde noch vorgeschrieben, daß bei der Benutzung von (10) die Größen t und h stets zwischen den Grenzen $\pm T$ und $\pm H$ bleiben sollen, so daß bezüglich der Konvergenz keinerlei Zweifel eintreten kann. Dies festgestellt werde nun die Funktion $f(x)$ von dem Argument $x = a$

ausgehend nach vorwärts und rückwärts mit dem Intervall h zwischen den Argumentgrenzen $a \pm Th$ tabuliert. Die zugehörige Haupttabelle enthält dann in der Spalte nullter Ordnung die Glieder

$$(a+p, 0) = f(a+ph),$$

wo p von $-T$ bis $+T$ läuft. Daraus ergibt sich nach (10) eine Reihendarstellung von der Gestalt

$$(a+p, 0) = A_0 f_0 + A_1 f_1 h + A_2 f_2 h^2 + \cdots, \tag{11}$$

worin die Koeffizienten A bestimmte rationale Zahlen bedeuten, deren Werte von der Form der tabulierten Funktion und ebenso von dem Intervall h unabhängig sind.

Berechnet man jetzt für die vorliegende Tafel aus (10) die Haupt- und Zwischenwerte des Differenzenschemas, so bleibt dabei die allgemeine Form der Reihe (11) erhalten, nur daß die Zahlenwerte der Koeffizienten A andere werden. Man darf also für die Differenzen allgemein ansetzen:

$$(a+p, q) = B_0 f_0 + B_1 f_1 h + B_2 f_2 h^2 + \cdots, \tag{12}$$

sodaß jetzt nur noch die Bestimmung der Koeffizienten B erforderlich ist, deren Werte allein von ihrer Stellennummer und den Zahlen p, q abhängen. Zu dem Ende ersetzen wir für den Augenblick die Funktion $f(x)$ durch die Exponentialfunktion $\exp x$, sodaß nach (8) und (9)

$$(a+p, q) = \exp(a)\ \exp(ph)\, S^q, \qquad (2p-q \text{ gerade})$$

$$(a+p, q) = \exp(a)\ \exp(ph)\, CS^q, \qquad (2p-q \text{ ungerade})$$

wird. Führt man die Ausdrücke in (12) ein und beachtet, daß die Größen $f_0, f_1, \cdots$ jetzt sämtlich gleich $\exp a$ sind, so entstehen für die Hauptwerte die Gleichungen

$$\exp(ph)\, S^q = B_0 + B_1 h + B_2 h^2 + \cdots, \tag{13}$$

für die Zwischenwerte dagegen die Gleichungen

$$\exp(ph)\, CS^q = B_0 + B_1 h + B_2 h^2 + \cdots. \tag{14}$$

Da die Entwicklung nach Potenzen von h bei den Größen C und S mit den Gliedern 1 und h anfängt, so beginnt die Entwicklung der linken Seiten von (13) und (14) mit dem Gliede h^q. Daraus fließen sofort die Gleichungen

$$B_0 = B_1 = \cdots = B_{q-1} = 0, \; B_q = 1.$$

Folglich wird wegen (12) in dem Differenzenschema der ursprünglichen Funktion $f(x)$

$$(a+p, q) = f_q h^q + B_{q+1} f_{q+1} h^{q+1} + \cdots. \tag{15}$$

Hiernach sind die q^{ten} Differenzen in Bezug auf das Tafelintervall h von der Ordnung q.

Ist das Intervall h so klein gewählt, daß in der Reihe (15) das Anfangsglied dem numerischen Betrage nach überwiegt, so sind die numerischen Größenordnungen der Differenzen $(a+p, 1)$, $(a+p, 2)$, $\cdot\cdot$ wesentlich durch die Ausdrücke

$$f_1 h, \quad f_2 h^2, \quad f_3 h^3, \cdots \tag{16}$$

gegeben. Hierdurch ist die Ursache aufgedeckt, welche bewirkt, daß bei genügend kleinem Intervall die Glieder des Differenzenschemas einer tabulierten Funktion den oben bei der Besprechung der Tabelle IV hervorgehobenen Gang zeigen.

Wird das Tafelintervall n-mal kleiner gewählt, so treten an die Stelle der Reihe (16) die Glieder

$$f_1 h : n, \quad f_2 h^2 : n^2, \quad f_3 h^3 : n^3, \cdots.$$

Man hat es also in der Gewalt, durch Verkleinerung des Intervalls die Differenzen einer höheren Ordnung bis zur Unmerklichkeit herabzudrücken. Als Beispiel möge das nachstehende Täfelchen dienen, welches der Tabelle IV gegenüber halbiertes Intervall besitzt.

x	Log x	I	II	III
50.0	1.6989700			
		+ 43214		
50.5	1.7032914		— 426	
		+ 42788		+ 8
51.0	1.7075702		— 418	
		+ 42370		+ 9
51.5	1.7118072		— 409	
		+ 41961		+ 8
52.0	1.7160033		— 401	
		+ 41560		+ 7
52.5	1.7201593		— 394	
		+ 41166		+ 7
53.0	1.7242759		— 387	
		+ 40779		
53.5	1.7283538			

Wie die Vergleichung mit Tabelle IV lehrt, stehen die Differenzen des neuen Täfelchens zu denen des alten — rund gerechnet — in dem Verhältnis 1 : 2, 1 : 4, 1 : 8. Wollte man in Tabelle IV das Intervall soweit verkleinern, daß die zweiten Differenzen die halbe Einheit der siebenten Decimale nicht übersteigen, so müßte der Verkleinerungsdivisor n aus der Gleichung

$$0.5 \quad \text{rund gleich} \quad 1600 : n^2$$

bestimmt werden, womit n rund gleich 57 werden würde.

§ 19. Um der Konvergenz der benutzten Reihen sicher zu sein, war bei der vorstehenden Beweisführung die vertikale Ausdehnung des Differenzenschemas auf $2T + 1$ Funktionswerte beschränkt worden, sodaß die Differenzen selber nur bis Ordnung $2T$ gehen. Der Satz von der numerischen Größenabstufung der Differenzen längs einer Zeile gilt daher zunächst auch nur bis zu dieser Ordnung $2T$ hin. Diese Beschränkung hat ihren guten Sinn. Wenn nämlich $f(x)$ für die Umgebung der Stelle $x = a$ nach dem *Taylor*schen Lehrsatze entwickelbar ist, so läßt sich durch Wahl eines hinreichend kleinen Intervalls allerdings erreichen, daß bei einer bestimmten im voraus festgesetzten Ordnungsnummer q die Differenzen der Argumentzeile a numerisch unmerklich werden. Dadurch ist aber keineswegs verbürgt, daß die späteren Differenzen jener Zeile nicht wieder merkliche, unter Umständen sogar unendlich große, Werte annehmen könnten. Setzt man z. B. für $f(x)$ den reziproken Wert von x und tabuliert die Funktion $1 : x$ mit den ganzzahligen Vielfachen von h als Argument, so enthält die Tafel auch das Argument $x = 0$ mit dem Funktionswerte ∞. Diese Unstetigkeit breitet sich nun in dem Differenzenschema nach oben und unten hin in derselben Weise aus, wie wir das in Tabelle III bei der Ausbreitung eines Fehlers kennen gelernt haben. Infolgedessen werden auf jeder Zeile des Schemas die Differenzen von einer gewissen Stelle ab unendlich.

Hiernach wird man im allgemeinen nur sagen dürfen, daß die numerische Abnahme, welche bei passend gewähltem Intervall die Differenzen einer Zeile anfangs zeigen, weiterhin je nach der Beschaffenheit der tabulierten Funktion andauern oder aber in Zunahme übergehen kann.

§ 20. Eine besondere Erwähnung verdient der Fall, wo $f(x)$ eine ganze rationale Funktion, sagen wir von der Ordnung s, ist. In diesem Falle verschwinden alle Ableitungen oberhalb der Ordnung s, folglich sind nach (15) die Größen $(a + p, q)$ beständig Null, sobald q oberhalb s liegt. Dem entsprechend besitzen die Differenzen der Ordnung s durch die ganze Tafel hindurch einen und denselben Wert, nämlich $f_s h^s$.

Der vorstehende Satz läßt sich zuweilen mit Vorteil bei der Tabulierung ganzer Funktionen verwenden. Es möge für den Augenblick das Differenzenschema der zu tabulierenden Funktion wieder wie in Tabelle I in der Gestalt

$$\begin{array}{cccccc} a_1 & & & & \\ & b_1 & & & \\ a_2 & & c_1 & & \\ & b_2 & & d_1 & \\ a_3 & & c_2 & & \\ . & . & . & . & . & . \end{array}$$

angesetzt werden. Ferner sei $f(x)$ eine ganze Funktion vom dritten Grade, sodaß die Größen d durchweg denselben Wert d_1 besitzen. Dann läßt sich, wenn auch noch die Anfangsglieder a_1, b_1, c_1 bekannt sind, die Tafel durch bloße Additionen berechnen, denn man hat nur nötig, zu der gegebenen Spalte der Größen d die drei nächsten Summenspalten mit den Anfangsgliedern c_1, b_1, a_1 zu bilden. Hierbei ist es zweckmäßig, die einzelnen Größen so anzuordnen, daß die jedesmal zu summierenden Zahlen unmittelbar untereinander stehen. Zu dem Ende denken wir uns das Schema, in das die Zahlen einzutragen sind, zunächst in der nachstehenden Weise angelegt.

Tabelle VII.

d_1	d_2	d_3	—
c_1	c_2	c_3	—
b_1	b_2	b_3	—
a_1	a_2	a_3	—

Bei der Ausführung berechnet man die vier ersten Funktionswerte der verlangten Tafel und bildet die sechs Differenzen dazu, sodaß a_1, b_1, c_1 und d_1 bekannt sind. Dann trägt man in die erste Zeile des Rechenschemas die Zahlen d ein, die der Voraussetzung nach sämtlich denselben Wert d_1 besitzen. Ferner trägt man c_1 ein und berechnet die übrigen c nach den Gleichungen

$$d_1 + c_1 = c_2, \quad d_2 + c_2 = c_3, \quad \text{u. s. w.} \tag{17}$$

Des weitern wird b_1 eingetragen und nach den Gleichungen

$$c_1 + b_1 = b_2, \quad c_2 + b_2 = b_3, \quad \text{u. s. w.}$$

gerechnet. Endlich findet man die Funktionswerte a, indem man nach Eintragung von a_1 die Gleichungen

$$b_1 + a_1 = a_2, \quad b_2 + a_2 = a_3, \quad \text{u. s. w.}$$

durchrechnet.

Als numerisches Beispiel diene die Berechnung der Kuben der natürlichen Zahlen. Der Anfang der Tafel hat in diesem Falle die Gestalt

0			
	1		
1		6	
	7		6
8		12	
	19		
27			

sodaß die konstante dritte Differenz gleich 6 wird, während

$$a_1 = 0, \quad b_1 = 1, \quad c_1 = 6$$

ist. Damit stellt sich die Rechnung folgendermaßen

6	6	6	6	—	—	—	—
6	12	18	24	30	—	—	—
1	7	19	37	61	91	—	—
0	1	8	27	64	125	216	—

Bei der Beschreibung des Rechnungsganges ist vorstehend der Raumersparnis halber die Anordnung der Tabelle VII zu Grunde gelegt worden, in welcher die Zahlen mit vierzeiligen Spalten nebeneinander angeordnet sind. Diese Anordnung ist an sich durchaus brauchbar, noch besser ist es jedoch, jene kurzen Spalten nicht nebeneinander, sondern untereinander zu setzen, weil dadurch die Kontrolle erleichtert wird. Trägt man nämlich in das solchergestalt abgeänderte Rechenschema zunächst die Größen d nebst c_1 ein und summiert diese Zahlen in einem Zuge, so entsteht sogleich das letzte c, das an der gehörigen Stelle einzutragen ist. Berechnet man darauf die übrigen c nach (17), so schließt die Rechnung mit dem bereits vorhandenen letzten c, wodurch die Summation kontrolliert ist. In derselben Weise wird bei der Ermittlung der Größen b und a verfahren.

Es bedarf keiner besonderen Auseinandersetzung, wie die Rechnung anzulegen ist, wenn statt der dritten Differenzenspalte eine andere Spalte des Schemas als Ausgangspunkt zu dienen hat.

§ 21. Das hier dargelegte Verfahren, das an Bequemlichkeit nichts zu wünschen übrig läßt, ist leider nur in beschränktem Maße anwendbar, sobald die als Ausgangspunkt der Rechnung dienenden Zahlen mit Abrundungsfehlern behaftet sind, weil die Wirkung solcher Fehler mit dem Fortschreiten der Summationen anwächst. Um hiervon eine deutliche Vorstellung zu gewinnen, benutzen wir den Satz, daß die Fehler eines Differenzenschemas zugleich das Differenzenschema der Fehler bilden. Es genügt also in Tabelle VII, die wiederum als Beispiel dienen möge, die Zahlen durch ihre Fehler zu ersetzen. Nimmt man der Einfachheit halber zunächst an, daß von den vier Größen a_1, b_1, c_1, d_1 nur d_1 fehlerhaft sei, und legt man ferner den Fehler von d_1 als Größeneinheit zu Grunde, so gestaltet sich die Rechnung für die Fehler, der Tabelle VII entsprechend, folgendermaßen:

1	1	1	1	1	—	—	—	—
0	1	2	3	4	5	—	—	—
0	0	1	3	6	10	15	—	—
0	0	0	1	4	10	20	35	—

Die vorstehenden Zahlen genügen um erkennen zu lassen, wie rasch

sich die Fehlerwirkung steigert, wenn man einerseits mit der Ordnungsnummer der Summenreihen, andererseits mit der Ausdehnung der Tafel vorwärts geht. Wenn also die Benutzung völlig scharfer Werte aus irgend welchen Gründen untunlich ist, so hat man bei der Anwendung des beschriebenen Summenverfahrens die Anzahl der mitzunehmenden Decimalen und die Ausdehnung der Operation derart zu bemessen, daß der Einfluß der Abrundungsfehler unschädlich bleibt. Hierbei ist die Bemerkung zu beachten, daß die Zeilen des obigen Täfelchens nichts anderes sind, als die Binomialkoeffizienten, die aus der Entwicklung der Ausdrücke

$$1 : (1 - t), \quad 1 : (1 - t)^2, \quad 1 : (1 - t)^3, \quad 1 : (1 - t)^4, \cdots$$

entspringen.

Bei der Tabulierung durch bloßes Summieren kann eine Additionsmaschine, namentlich wenn sie mit einer Druckvorrichtung versehen ist, häufig gute Dienste leisten. Es ist auch unschwer zu übersehen, daß es möglich sein muß, durch Kuppelung mehrerer solcher Werke eine Maschine herzustellen, die z. B. eine Funktion vierten Grades selbsttätig tabuliert, sobald die Anfangsglieder der Spalten des Differenzenschemas in der Maschine eingestellt worden sind. Die nur in wenigen Exemplaren vorhandenen Maschinen, die von *Scheutz* Vater und Sohn, sowie von *Wijkander* konstruiert worden sind, scheinen — nach den spärlichen darüber veröffentlichten Notizen zu schließen — in der Tat auf diesem Grundgedanken zu beruhen.

§ 22. Der Satz, daß bei hinreichend kleinem Intervall die Differenzen merklich regelmäßig laufen müssen, enthält ein wertvolles Mittel zur Kontrolle der Tabulierungsrechnung, denn der bei richtiger Rechnung regelmäßige Gang der Differenzen wird durch einen in den Funktionswerten begangenen Fehler notwendig gestört. Nach dem, was in § 14 über die Ausbreitung eines Fehlers gesagt worden ist, wird man, wenn die letzte Differenzenreihe stark springt, den Verdacht zunächst auf denjenigen Funktionswert richten, auf dessen Zeile der stärkste Sprung stattfindet.

Die Kontrolle durch Differenzen vermag selbstverständlich nicht die durchlaufenden Fehler aufzudecken, wie sie z. B. aus der Benutzung einer falschen Formel oder falscher Konstanten entspringen. Ferner ist zu beachten, daß die Funktionswerte im allgemeinen mit unvermeidlichen Abrundungsfehlern behaftet sind, deren obere Grenze auch bei völlig korrekter Rechnung nicht unter die halbe Einheit der letzten Decimale herabgedrückt werden kann. Diese Abrundungsfehler machen sich natürlich in den Differenzen geltend und erzeugen Sprünge, die um so merklicher werden, je höher die Ordnungsnummer der Differenz ist. Als Erläuterung möge das folgende mit acht

Stellen angesetzte Täfelchen dienen, in welchem die achte Stelle der besseren Übersicht halber durch einen Punkt abgetrennt ist.

x	Log x	I	II	III	IV
50	1.6989700·0				
		+ 86001·8			
51	1.7075701·8		— 1670·2		
		+ 84331·6		+ 63·9	
52	1.7160033·4		— 1606·3		— 4·0
		+ 82725·3		+ 59·9	
53	1.7242758·7		— 1546·4		— 3·1
		+ 81178·9		+ 56·8	
54	1.7323937·6		— 1489·6		— 3·1
		+ 79689·3		+ 53·7	
55	1.7403626·9		— 1435·9		
		+ 78253·4			
56	1.7481880·3				

Vergleicht man die vorstehenden Zahlen mit Tabelle IV, welche dieselben Logarithmen siebenstellig enthält, so erkennt man deutlich, wie die kleinen Sprünge in den vierten Differenzen der Tabelle IV durch die Abrundung erzeugt worden sind.

Setzt man in der Tabelle III, welche die Ausbreitung eines einzelnen Fehlers darstellt, für x den Betrag 0.5, wobei als Einheit die letzte Decimale der Rechnung gedacht ist, so ergeben sich für die einzelnen Ordnungsnummern q folgende Maximalwerte der Wirkung eines einzelnen Fehlers:

$q =$	0	1	2	3	4	5	6
Fehler =	0.5	0.5	1.0	1.5	3.0	5.0	10.0.

Die halbe Einheit im Funktionswerte wirkt also in der sechsten Differenz schon mit zehn Einheiten. Dieser Umstand ist zu beachten, wenn man die Sprünge in den höheren Differenzen richtig beurteilen will.

§ 23. Zum Schlusse dieses Abschnitts möge, anknüpfend an die Bemerkungen über Abrundungsfehler, noch kurz der Fall der zweifelhaften Abrundung berührt werden. Nach einer bekannten Vorschrift wird, um den Abrundungsfehler möglichst herunterzudrücken, die letzte mitzunehmende Stelle unverändert gelassen oder aber um eine Einheit erhöht, je nachdem der Betrag der abzuwerfenden Decimalen unter oder über der halben Einheit jener Stelle liegt. Diese Vorschrift versagt jedoch, wenn der abzuwerfende Betrag genau eine halbe Einheit der letzten Stelle ausmacht: ob z. B. 13.5 auf 13 oder auf 14 abzurunden ist, das läßt sich nach der genannten Vorschrift nicht entscheiden. Damit wird man zu der Frage geführt, ob in solchen Fällen der Rechner einfach nach Gutdünken verfahren oder

irgend eine feste Regel befolgen oder endlich die 5 einfach stehen lassen und damit weiter rechnen solle. Der letzte Weg erscheint zunächst als der beste, denn er schneidet jeden Zweifel ab und bewirkt zugleich eine kleine Erhöhung der Genauigkeit der Rechnung — er ist aber wegen anderer Gründe nicht immer gangbar. Um das deutlich zu machen, betrachten wir einen Fall, der bei der Bearbeitung beobachteter Zahlen gar nicht selten ist.

In einer Reihe von Winkelmessungen wird ein geteilter Kreis immer an je zwei gegenüberstehenden Strichen abgelesen; diese Ablesungspaare werden bis auf ganze Sekunden notiert und sind vor der weiteren Benutzung zu mitteln; die Genauigkeit der Beobachtungen ist so beschaffen, daß die Notierung der ganzen Sekunden eine mehr als ausreichende Schärfe gewährt, daß also die Ansetzung der Zehntelsekunden einen völlig überflüssigen Ziffernballast bedeutet, der von vornherein vermieden werden soll, weil jene Mittelwerte samt den weiteren Resultaten nachher durch den Druck zu veröffentlichen sind; wie hat sich der Rechner beim Abwerfen der 5 zu verhalten, die ja durchschnittlich in fünfzig Fällen unter hundert auftreten wird? Es kann nicht zweifelhaft sein, daß unter den angeführten Umständen die Befolgung einer an sich zulässigen festen Regel besser ist, als Willkür, namentlich dann, wenn das vorliegende Beobachtungsmaterial der Kontrolle halber von verschiedenen Rechnern doppelt durchgerechnet wird. Gewöhnlich wird für diese Fälle vorgeschlagen, auf „gerade“ abzurunden, d. h. für die Beträge

$$0.5, \quad 2.5, \quad 4.5, \quad 6.5, \quad 8.5$$

die Zahlen 0, 2, 4, 6, 8 und für die Beträge

$$1.5, \quad 3.5, \quad 5.5, \quad 7.5, \quad 9.5$$

die Zahlen 2, 4, 6, 8, 10 anzusetzen. Man erreicht dabei zugleich, daß die Fehler bald positiv, bald negativ ausfallen, daß also innerhalb einer längeren zusammenhängenden Rechnung im allgemeinen eine gewisse Ausgleichung zu erwarten ist. Der Vorschlag ist zweifellos annehmbar, man hat sich nur gegenwärtig zu halten, daß er nicht der einzige brauchbare ist, und daß unter Umständen eine andere Vorschrift zweckmäßiger sein kann.

Zweiter Abschnitt.

Interpolation bei Tafeln.

§ 24. Das Wort „Interpolieren" besitzt ursprünglich die Bedeutung des auf Täuschung berechneten Zustutzens einer Sache. Der mathematische Sprachgebrauch sieht von dieser Bedeutung ab und verwendet das Wort in einem anderen, teils engeren, teils weiteren Sinne: im engeren Sinne bezeichnet es die ergänzende Einschaltung bei Tafeln, im weiteren dagegen die genäherte Darstellung von Gebilden oder Vorgängen durch geeignete Formeln. Hier haben wir es zunächst nur mit der Interpolation von Tafeln zu tun. Die zu lösende Aufgabe kommt dann, allgemein gefaßt, auf folgendes hinaus: Für gewisse vorgeschriebene Werte des Arguments x sind die Werte einer vorgelegten Funktion $f(x)$ bekannt; gesucht wird der Wert von $f(x)$ für irgend ein anderes Argument. Oder geometrisch gesprochen: von einer Kurve sind gewisse Punkte bekannt; gesucht werden weitere Punkte der Kurve.

In der vorstehenden Fassung kann die Aufgabe, wie man sofort erkennt, je nach Umständen bestimmt oder unbestimmt oder überbestimmt sein. Daher sind, wenn eine bestimmte Lösung zu stande kommen soll, gewöhnlich noch gewisse Nebenbedingungen hinzuzufügen. In welcher Weise das geschieht, wird sich weiterhin zeigen.

Die Interpolationsformeln, welche seither bei den Rechnern in Gebrauch gekommen sind, stammen aus einer gemeinsamen Quelle her, nämlich aus einer Formel, die zuerst von *Waring (Problems concerning interpolations, Philosophical Transactions of the R. Society of London, Vol. 69, 1779)* abgeleitet worden ist, die jedoch von alters her unter dem Namen von *Lagrange* geht. Diese Formel soll zunächst entwickelt werden.

§ 25. Es sei die ganze Funktion n^{ten} Grades

$$y = f(x) = a_0 + a_1 x + \cdots + a_n x^n \tag{1}$$

vorgelegt, deren geometrische Darstellung mit den rechtwinkligen Koordinaten x und y auf eine sogenannte Parabel n^{ter} Ordnung führt.

Von dieser Parabel seien $n + 1$ Punkte dadurch gegeben, daß man zu den irgendwie vorgeschriebenen Abscissen

$$x_0, x_1, \cdots x_n$$

die Ordinaten

$$y_0 = f(x_0), \quad y_1 = f(x_1), \cdots y_n = f(x_n)$$

kennt. Sollen nun weitere Punkte gefunden werden, so kann man zunächst in *der* Weise vorgehen, daß man für jeden gegebenen Punkt $(x_p y_p)$ nach (1) die Gleichung

$$y_p = a_0 + a_1 x_p + \cdots a_n x_p^n \tag{2}$$

ansetzt, darauf das System der so entstehenden $n + 1$ Gleichungen nach den Koeffizienten a auflöst und die gefundenen Werte in (1) einsetzt. Dieser Rechnungsgang führt zu einer völlig bestimmten Lösung, sobald die vorgeschriebenen Abscissen x_p, wie hier selbstverständlich ist, voneinander verschieden sind; er führt überdies zur vollständigen Kenntnis der Funktion $f(x)$, ist indessen trotzdem nur dann zu empfehlen, wenn man aus irgend welchen Gründen die Kenntnis der Koeffizienten nötig hat. Wird dagegen nur verlangt, daß man zu den gegebenen Punkten $(x_p y_p)$ noch einige neue hinzufügt, so ist es vorzuziehen, die Koeffizienten gar nicht erst zu bestimmen, sondern sie von vornherein aus den Gleichungen (1) und (2) zu eliminieren, wobei folgender Weg eingeschlagen werden soll.

Mit der willkürlich gewählten Konstante C und den gegebenen Abscissen bilden wir das Produkt

$$T(x) = C(x - x_0)(x - x_1) \cdots (x - x_n), \tag{3}$$

das entwickelt ein Polynom von der Ordnung $n + 1$ liefert. Entwickelt man $T(x)$ nach Potenzen von $x - x_p$ und bezeichnet die Ableitungen nach x durch Accente, so erhält man, da $T(x_p)$ verschwindet, die Reihe

$$T(x) = T'(x_p)(x - x_p) + \tfrac{1}{2} T''(x_p)(x - x_p)^2 + \cdots.$$

Dividiert man beiderseits durch das erste Glied der rechten Seite und setzt

$$U(x, x_p) = T(x) : T'(x_p)(x - x_p), \tag{4}$$

so ist $U(x, x_p)$ eine ganze Funktion n^{ter} Ordnung von x, die für $x = x_p$ den Wert Eins annimmt, für die übrigen gegebenen Abscissen dagegen verschwindet. Infolgedessen nimmt das Aggregat

$$g(x) = y_0 U(x, x_0) + y_1 U(x, x_1) + \cdots + y_n U(x, x_n)$$

für $x = x_p$ den Wert y_p an. Hiernach verschwindet der Ausdruck

$$h(x) = f(x) - g(x),$$

der eine ganze Funktion höchstens n^{ter} Ordnung darstellt, für die $n+1$ gegebenen Abscissen x_p; er müßte also, falls er nicht identisch null ist, die $n+1$ Teiler $x-x_p$ besitzen. Da das nicht möglich ist, so verschwindet $h(x)$ identisch, folglich ist $g(x)$ identisch gleich $f(x)$. Damit erhält man die sogenannte *Lagrange*sche *Interpolationsformel* in der Gestalt

$$f(x) = \sum_p y_p U(x, x_p) = \sum_p f(x_p)\, U(x, x_p), \qquad (p = 0, 1, \cdots n). \quad (5)$$

Nach (4) ist

$$U(x, x_p)\, T'(x_p) = T(x) : (x - x_p).$$

Da hierin für $x = x_p$ der Ausdruck $U(x_p, x_p)$ gleich Eins wird, so ist $T'(x_p)$ gleich dem Werte, den der Quotient rechterhand für $x = x_p$ annimmt. Dieser Quotient wird aber erhalten, wenn man in dem Produkt (3) den Faktor $x - x_p$ streicht. Folglich darf man, da der Faktor C sich heraushebt, schreiben

$$U(x, x_p) = \prod_q \frac{x - x_q}{x_p - x_q}, \tag{6}$$

wo das Produkt über die Nummern $q = 0, 1, \cdots n$ zu nehmen ist, jedoch unter Ausschluß der Nummer $q = p$.

Die Formeln (5) und (6), die namentlich für die logarithmische Rechnung bequem sind, lassen mancherlei Umformungen zu, von denen aus man durch geeignete Spezialisierung zu den Interpolationsformeln gelangen kann, welche bei dem Gebrauch der numerischen Tafeln in Betracht kommen. Es ist jedoch zweckmäßiger, die gesuchten Formeln direkt abzuleiten und erst nachher ihren Zusammenhang mit der Darstellung von *Lagrange* nachzuweisen.

Da es bei der Herleitung von (5) auf die Werte der Koeffizienten in dem betrachteten Polynom

$$f(x) = a_0 + a_1 x + \cdots + a_n x^n$$

nicht weiter ankommt, so darf man sich $f(x)$ z. B. auf das Anfangsglied a_0 reduziert denken. In diesem Falle geht (5) in die Identität

$$1 = \sum_p U(x, x_p)$$

über. Hiernach läßt sich der in (5) gegebene Ausdruck als ein Mittelwert ansehen, der mit den Gewichten $U(x, x_p)$ aus den Größen $f(x_p)$ gebildet worden ist, wobei jedoch zu beachten ist, daß die Größen U nicht sämtlich positiv zu sein brauchen.

§ 26. Es sei $f(x)$ wieder eine ganze Funktion n^{ter} Ordnung von x und es möge $f(a+th)$ nach t entwickelt den Ausdruck

$$f(a + th) = F_0 + F_1 th + \cdots + F_n (th)^n \tag{7}$$

liefern. Wir denken uns nun $f(x)$ von der Stelle $x = a$ aus nach vorwärts und rückwärts mit dem Intervall h tabuliert, was darauf hinauskommt, daß man $(f a + t h)$ als Funktion von t betrachtet und für die ganzzahligen Werte von t tabuliert. Weiter bilde man aus den Funktionswerten das Differenzenschema mit seinen Haupt- und Zwischengrößen, die wieder durch das Symbol $(a + p, q)$ bezeichnet werden sollen, wie das in § 16 beschrieben worden ist.

Um die aus (7) folgenden Ausdrücke für die einzelnen Differenzen zu erhalten, kann man nach den Bemerkungen in § 13 zunächst die Haupt- und Zwischenwerte zu jedem einzelnen Gliede der Reihe (7) aufstellen und dann die so erhaltenen Schemas additiv verbinden. Nun ist die Spalte o^ter^ Ordnung in dem Schema für die Potenz t^r nichts anderes als eine Tafel der r^{ten} Potenzen der natürlichen Zahlen, folglich sind die zugehörigen Differenzen in der Haupttabelle ganze Zahlen und in der Zwischentabelle ganze Vielfache von $\frac{1}{2}$. Bezeichnet man die dem Symbol $(a + p, q)$ entsprechende Differenz von t^r mit (r, p, q) und beachtet die in § 18 gegebene Formel (15), so erhält man, weil die r^{te} Ableitung von t^r gleich $r!$ ist, für die r^{ten} Differenzen von t^r den Ausdruck

$$(r, p, r) = r! \tag{8}$$

Ferner verschwinden die Differenzen oberhalb der Ordnung r, folglich ist

$$(r, p, q) = 0 \quad \text{für } q > r. \tag{9}$$

Wendet man diese Bemerkungen auf den in (7) aufgestellten Ausdruck

$$f(a + th) = \sum_r F_r (th)^r, \quad (r = 0, 1, \cdot \cdot n) \tag{10}$$

an, so erhält man die zu $f(x)$ gehörigen Differenzen in der Gestalt

$$(a + p, q) = \sum_r (r, p, q) F_r h^r.$$

Bei der weiteren Verwendung der vorstehenden Gleichungen denken wir uns aus jeder Differenzenspalte von $f(x)$ je ein Glied herausgegriffen und bezeichnen die Zeilennummer des ausgewählten Gliedes q^{ter} Ordnung mit $a + p_q$. Dann entspricht, wenn man q von 0 bis n laufen läßt, den ausgewählten Gliedern das Gleichungssystem

$$\left.\begin{aligned}
(a + p_0, 0) &= \sum_r (r, p_0, 0) F_r h^r, \\
(a + p_1, 1) &= \sum_r (r, p_1, 1) F_r h^r, \\
&\cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \\
(a + p_n, n) &= \sum_r (r, p_n, n) F_r h^r.
\end{aligned}\right\} \tag{11}$$

Die vorstehenden $n + 1$ Gleichungen verbinde man jetzt mit den noch näher zu bestimmenden Multiplikatoren $A_0, A_1, \cdot \cdot A_n,$ ziehe die Ver-

bindung von (10) ab und bestimme nunmehr die Größen A_r so, daß die Koeffizienten $F_0, F_1, \cdots$ herausfallen. Dann entsteht erstens die Gleichung

$$f(a + th) = \sum_s A_s(a + p_s, s), \quad (s = 0, 1, \cdots n) \tag{12}$$

und zweitens das System der Bedingungen

$$t^r = \sum_s (r, p_s, s) A_s. \tag{13}$$

Schreibt man die vorstehenden Bedingungen aus und beachtet dabei die Beziehungen (8) und (9), so wird für $r = 0, 1, \cdots$

$$\left.\begin{aligned} 1 &= 0!\, A_0, \\ t &= (1, p_0, 0) A_0 + 1!\, A_1, \\ t^2 &= (2, p_0, 0) A_0 + (2, p_1, 1) A_1 + 2!\, A_2, \\ t^3 &= (3, p_0, 0) A_0 + (3, p_1, 1) A_1 + (3, p_2, 2) A_2 + 3!\, A_3, \\ &\cdots\cdots\cdots\cdots\cdots\cdots \end{aligned}\right\} \tag{14}$$

Hiernach ist A_q eine jederzeit völlig bestimmte ganze Funktion q^{ter} Ordnung von t, die im übrigen nur abhängt von dem Index q und den ausgewählten Zeilennummern $p_0, p_1, \cdots p_{q-1}$, dagegen unabhängig ist von den folgenden Zeilennummern und von der Ordnungszahl n der ganzen Funktion $f(x)$.

Die Gleichung (12) liefert offenbar für die ganze Funktion $f(x)$ eine neue Darstellung, die sich von der *Lagrange*schen Formel dadurch unterscheidet, daß an Stelle der gegebenen Funktionswerte gewisse Glieder des Differenzenschemas als bekannt vorausgesetzt werden.

§ 27. Die Gleichung (12), die fortan kurz als die „*Formel J*" bezeichnet werden soll, löst, wenn $f(x)$ eine ganze Funktion höchstens n^{ter} Ordnung bedeutet, in aller Strenge die Aufgabe, zwischen die gegebenen Kurvenpunkte beliebige neue Punkte einzuschalten. Wird dagegen irgend eine andere tabulierte Funktion nach der Formel J behandelt, so erhält man im allgemeinen nicht den wahren Wert von $f(a + th)$. Setzt man

$$K = f(a + th) - \sum_s A_s(a + p_s, s), \tag{15}$$

so gibt K offenbar die Korrektion an, die man an das Rechnungsergebnis der Formel J anzubringen hat, um den richtigen Wert von $f(a + th)$ zu erhalten. Diese Korrektion ist also bei einer ganzen Funktion n^{ter} Ordnung identisch null, sonst aber im allgemeinen von Null verschieden.

Um das Verhalten von K zu untersuchen, nehmen wir an, daß $f(x)$ die *Taylor*sche Entwicklung in der Gestalt

$$f(a + th) = F_0 + F_1 th + \cdots + F_n t^n h^n + R(x) \tag{16}$$

zulasse, wo $R(x)$ den sogenannten *Rest* der Entwicklung bedeutet. Weiter denke man sich $R(x)$ in derselben Weise wie $f(x)$ tabuliert, wobei die zu $f(x)$ gehörigen Differenzen wie bisher mit $(a+p, q)$, die entsprechenden Differenzen von $R(x)$ dagegen mit $R(p, q)$ bezeichnet werden sollen. Wird dann die Darstellung (16) in den Ausdruck für K eingesetzt, so heben sich die Potenzglieder, die ja nur bis zur n^{ten} Ordnung aufsteigen, heraus, und man erhält

$$K = R(a+th) - \sum_s A_s R(p_s, s). \tag{17}$$

Hiernach hängt der Verlauf der Korrektion K wesentlich von dem Verhalten des Restes $R(x)$ ab.

Für den Rest $R(x)$ wird in den Lehrbüchern eine Integraldarstellung abgeleitet, die hier benutzt werden soll. Schreibt man die *Taylor*sche Reihe zunächst in der Gestalt

$$f(z) = f(x) + (z-x) f(x)_1 + \cdots + \frac{(z-x)^n}{n!} f(x)_n + R,$$

wo die Indices die Ableitungen anzeigen, so stellt sich R als eine Funktion von x und z dar. Die partielle Ableitung von R nach x nimmt, wenn man gehörig reduziert, die Gestalt

$$\frac{\partial R}{\partial x} = -\frac{(z-x)^n}{n!} f(x)_{n+1}$$

an. Integriert man wieder nach x und beachtet, daß R für $x = z$ verschwindet, so wird

$$R = -\int_z^x du \frac{(z-u)^n}{n!} f(u)_{n+1}.$$

Ersetzt man hierin x durch a, z durch $a + th$, u durch $a + th - vh$, wo v die neue Integrationsveränderliche bedeutet, so läßt sich das Restglied nunmehr durch die Formeln

$$R(x) = R'(x) h^{n+1}, \tag{18}$$

$$R'(x) = R'(a+th) = \frac{1}{n!}\int_0^t dv\, v^n f(a+th-vh)_{n+1} \tag{19}$$

ausdrücken. Denkt man sich für R' das Differenzenschema in derselben Weise wie für R gebildet und die Differenzen von R' ähnlich wie bei R durch die Symbole $R'(p, q)$ bezeichnet, so kann man für die Korrektion K statt (17) die Gleichungen

$$K = K' h^{n+1}, \tag{20}$$

$$K' = R'(a+th) - \sum_s A_s R'(p_s, s) \tag{21}$$

ansetzen. Hiernach handelt es sich jetzt um das Verhalten von R' und K'.

§ 28. Um bei der weiteren Untersuchung die Vorstellung zu fixieren, mögen die nachstehenden Bestimmungen getroffen werden. Das Argument a und das System der Nummern $p_0, p_1, \cdot \cdot$ werde als fest betrachtet, desgleichen auch die Ordnungszahl n der höchsten Differenz, die in der Formel J mitgenommen worden ist. Weiter möge bei den vorzunehmenden Tabulierungen das Differenzenschema stets so geschrieben werden, daß das Argument von oben nach unten wächst, daß also das Intervall h stets positiv ist. Die Größe h soll hierbei verschiedener Werte fähig sein, mit der Einschränkung, daß h stets unter einer passend zu wählenden festen Zahl H bleibt. Endlich werde, entsprechend der bei den Anwendungen bestehenden Übung, die Größe t in dem Argument $a + th$, für welches $f(x)$ zu interpolieren ist, auf echt gebrochene Werte beschränkt.

Um aus (19) die Differenzen $R'(p_s, s)$ zu finden, muß man in der Tafel für R' diejenigen Funktionswerte aufsuchen, von denen jene Differenzen abhängen. Die zu den genannten Funktionswerten gehörigen Argumente sind, sobald über das Intervall h verfügt worden ist, durch die Größe a und die Nummern p_s vollständig bestimmt und sollen für den Augenblick kurz als „Argumentreihe I“ bezeichnet werden. Ferner führt das Glied $R'(a + th)$ in (19) auf eine „Argumentreihe II“, wenn t alle zulässigen Werte annimmt. Die Reihen I und II zusammengenommen dehnen sich dann im Gebiete der Veränderlichen x über eine gewisse Strecke aus, die als Strecke $S(h)$ bezeichnet werden soll und von dem Argument $x = B(h)$ bis zu dem Argument $x = C(h)$ reichen möge. Die Abstände der Stellen $B(h)$ und $C(h)$ von der Stelle a sind der Größe h proportional, und das Gleiche gilt von der Ausdehnung der Strecke $S(h)$. Setzt man also

$$B(h) = a + hU, \quad C(h) = a + hV,$$

so sind U und V nur von den Nummern p_s abhängig, hingegen unabhängig von h; ferner wird, wenn $C(h)$ oberhalb $B(h)$ liegt,

$$S(h) = (U - V)h.$$

Dies festgestellt, gehen wir jetzt auf den Ausdruck von R' in (19) zurück. Die obere Integralgrenze t hat, wenn (21) zur Berechnung von K' benutzt wird, Beträge anzunehmen, die der Wertstrecke von $t = U$ bis $t = V$ angehören. Da ferner in (19) das Argument der unter dem Integralzeichen stehenden Ableitung f_{n+1} von a bis $x = a + th$ läuft, so kommen bei der Berechnung von K' die Werte von f_{n+1} in Betracht, deren Argument der Strecke $S(h)$ angehört.

Es werde nunmehr angenommen, daß die Funktion $f(x)_{n+1}$ dem absoluten Betrage nach unterhalb einer bestimmten angebbaren Größe g bleibe, solange sich x innerhalb der Strecke $S(H)$ bewegt, die zu dem größten für h zugelassenen Werte H gehört. Dann

bestehen, wenn man die Berechnung von K' nach (21) für alle zu lässigen Wertepaare der Größen t, h ausgeführt, für die dabei auftretenden absoluten Beträge der Größen R' und der Koeffizienten A_s bestimmte angebbare und von h unabhängige obere Grenzen, und daraus folgt dann weiter für den absoluten Betrag von K' eine bestimmte angebbare und von h unabhängige obere Grenze, die mit k bezeichnet werden möge. Bezeichnet also e einen echten Bruch, so läßt sich K' in der Gestalt ek ansetzen, woraus dann weiter

$$K = ekh^{n+1}$$

folgt. Hiernach kann die Korrektion K durch Einführung eines hinreichend kleinen Tafelintervalls beliebig klein gemacht werden.

Die über f_{n+1} gemachte Voraussetzung läßt sich erfüllen, wenn f_{n+1} für eine gewisse endliche Umgebung der Stelle $x = a$ endlich bleibt, denn man hat dann die Größe H nur derart festzusetzen, daß die zugehörige Strecke $S(H)$ ganz in jene Umgebung fällt. Demnach darf man folgenden Satz aussprechen: *Die Korrektion der mit $n + 1$ Gliedern für $f(a + th)$ angesetzten Interpolationsformel J läßt sich durch Einführung eines hinreichend kleinen Tafelintervalls numerisch unmerklich machen, wenn die Ableitung $f(x)_{n+1}$ für eine gewisse, zu beiden Seiten der Stelle $x = a$ endlich ausgedehnte Argumentstrecke endlich bleibt.*

Der vorstehende Satz beseitigt offenbar unter gewissen Voraussetzungen die Beschränkung auf ganze Funktionen, die der Formel J ursprünglich anhaftete.

§ 29. Die vorstehenden Betrachtungen können noch anders gewendet werden. Bisher war die Ordnungszahl n, von der die Gliedermenge in der Formel J abhängt, als fest angesehen worden, während das Tafelintervall h wechselnde und im besondern beliebig kleine Werte annehmen konnte. Im folgenden möge nun umgekehrt die Größe h festgehalten werden, während die Zahl n der Reihe nach die Werte 1, 2, 3, ·· annehmen soll. Dann kann man fragen, ob die Korrektion K, die man jedesmal der Formel J hinzuzufügen hat, mit unbeschränkt wachsendem n gegen Null geht oder nicht. Trennt man die verschiedenen Möglichkeiten, so lassen sich folgende Fälle unterscheiden: 1) K konvergiert gegen Null, 2) K konvergiert gegen einen von Null verschiedenen endlichen Wert, 3) K wächst nach einer und derselben Seite hin ins unendliche, 4) K schwankt hin und her.

Im ersten Falle ist die unendliche Reihe

$$A_0(a + p_0, 0) + A_1(a + p_1, 1) + \cdots \tag{22}$$

konvergent und besitzt die Summe $f(a + th)$. Im zweiten Falle ist die Reihe wiederum konvergent, führt aber auf eine von $f(a + th)$

verschiedene Summe. Im dritten und vierten Falle ist die Reihe divergent. In allen vier Fällen kann es vorkommen, daß K sogleich anfangs, also für mäßige Werte von n, auf einen für die Rechnung unmerklichen Betrag herabsinkt. Tritt dies ein, so ist offenbar der eigentliche Zweck der Formel J erfüllt: das spätere Verhalten der Korrektion K wird für den Rechner gleichgültig, weil er die späteren Glieder der Reihe überhaupt nicht mehr braucht. Im Falle der Divergenz hat man es dann mit demjenigen Reihentypus zu tun, welcher als *semikonvergent* bezeichnet wird.

Der Fall der Semikonvergenz ist als etwas sehr gewöhnliches anzusehen. Verfolgt man nämlich den Verlauf der vorgelegten Funktion für das ganze reelle Gebiet der Veränderlichen x, so kommt es oft genug vor, daß die Funktion an einer oder an mehreren Stellen unendlich oder unbestimmt oder imaginär wird. Solche „Irregularitäten" breiten sich nun ähnlich wie der Fehler eines Funktionswertes fächerförmig in dem Differenzenschema aus und können dann in der unendlichen Reihe (22) irreguläre Glieder erzeugen, welche die Reihe divergent machen. Ein einfaches Beispiel hierfür liefern die gewöhnlichen logarithmischen Tafeln.

Ist der gesuchte Funktionswert $f(a + th)$ nur wenige Intervalle von einer irregulären Stelle entfernt, so versagt im allgemeinen die Formel J, indem die Korrektion K überhaupt nicht oder doch nicht rasch genug auf einen unmerklichen Betrag herabsinkt. Man kann dann in einem solchen Teile der Tafel nicht interpolieren und hat die Berechnung von $f(x)$ abzuändern. Der gewöhnlich benutzte Ausweg besteht darin, daß man für die in Betracht kommende Gegend der Tafel an Stelle der Funktion $f(x)$ eine andere regulär verlaufende Funktion $g(x)$ tabuliert, die so gewählt ist, daß aus ihr $f(x)$ durch eine kleine Nebenrechnung gefunden werden kann. So werden z. B. für kleine Winkel neben den Werten von Log sin x und Log tg x Täfelchen gegeben, welche die Logarithmen der Quotienten sin $x : x$ und tg $x : x$ enthalten.

§ 30. Die Formel J liefert zunächst so viele besondere Interpolationsformeln, als man besondere Wertsysteme für die Zeilennummern $p_0, p_1, \cdots$ aufstellen kann. Bei der Auswahl dieser Nummern wird man selbstverständlich darauf ausgehen, daß die Rechnungsvorschrift möglichst übersichtlich und bequem ausfällt. Dem entsprechend sind seither nur die vier weiterhin behandelten Formeln bei den Rechnern in allgemeinen Gebrauch gekommen. Diese nach *Newton*, *Gauß*, *Stirling* und *Bessel* benannten Formeln sollen im folgenden kurz durch die Anfangsbuchstaben *N*, *G*, *S* und *B* bezeichnet werden; die jedesmal benutzten Glieder des Differenzenschemas sind aus der nachstehenden Übersicht zu entnehmen:

Formel N, *Newton.*

$$(a, 0) \quad (a + \tfrac{1}{2}, 1) \quad (a + 1, 2) \quad (a + \tfrac{3}{2}, 3) \quad (a + 2, 4) \cdots$$

Formel G, *Gauß.*

$$(a, 0) \quad (a + \tfrac{1}{2}, 1) \quad (a, 2) \quad (a + \tfrac{1}{2}, 3) \quad (a, 4) \quad \cdots$$

Formel S, *Stirling.*

$$(a, 0) \quad (a, 1) \quad (a, 2) \quad (a, 3) \quad (a, 4) \quad \cdots$$

Formel B, *Bessel.*

$$(a + \tfrac{1}{2}, 0) \quad (a + \tfrac{1}{2}, 1) \quad (a + \tfrac{1}{2}, 2) \quad (a + \tfrac{1}{2}, 3) \quad (a + \tfrac{1}{2}, 4) \cdots$$

Außer diesen Formeln sollen noch zwei mit G' und G'' bezeichnete Varianten der Formel G berücksichtigt werden, bei denen folgende Glieder des Differenzenschemas zur Verwendung kommen:

Formel G'.

$$(a, 0) \quad (a - \tfrac{1}{2}, 1) \quad (a, 2) \quad (a - \tfrac{1}{2}, 3) \quad (a, 4) \quad \cdots$$

Formel G''.

$$(a + 1, 0) \quad (a + \tfrac{1}{2}, 1) \quad (a + 1, 2) \quad (a + \tfrac{1}{2}, 3) \quad (a + 1, 4) \cdots$$

Sämtlichen Formeln ist das Merkmal gemeinsam, daß sich je zwei benachbarte Glieder in der Zeilennummer höchstens um eine halbe Einheit unterscheiden. Bei N bilden die Glieder in dem Differenzenschema die Stufen einer absteigenden Treppe, während sie bei G, G' und G'' einer Hauptzeile und der nachfolgenden oder vorausgehenden Zwischenzeile angehören, bei S und B dagegen auf einer einzigen durchgehenden Zeile stehen.

§ 31. Bei der Aufsuchung der Koeffizienten genügt es für $f(x)$ eine ganze Funktion n^{ter} Ordnung zu Grunde zu legen, sodaß die Formeln jedesmal mit der Differenz n^{ter} Ordnung abbrechen. Die Formel N erscheint dann in der Gestalt

$$f(a + th) = A_0(a, 0) + A_1(a + \tfrac{1}{2}, 1) + \cdots + A_n\left(a + \frac{n}{2}, n\right). \quad (23)$$

Die Funktionswerte, aus denen die benutzten Differenzen entspringen, sind durch die Größen

$$(a, 0), \quad (a + 1, 0), \quad \cdots (a + n, 0) \qquad (24)$$

gegeben, die jetzt kurz mit $y_0, y_1, \cdots y_n$ bezeichnet werden sollen. Es ist dann

$$\begin{aligned}
(a, 0) &= y_0,\\
(a + \tfrac{1}{2}, 1) &= y_1 - y_0,\\
(a + 1, 2) &= y_2 - 2y_1 + y_0,\\
(a + \tfrac{3}{2}, 3) &= y_3 - 3y_2 + 3y_1 - y_0.\\
&\cdots\cdots\cdots
\end{aligned}$$

Setzt man diese Ausdrücke in (23) ein, so erscheint $f(a+th)$ in der Gestalt

$$f(a+th) = y_0 U_0 + y_1 U_1 + \cdots + y_n U_n. \qquad (25)$$

Hierin sind die Größen U lineare Verbindungen der Koeffizienten A; im besondern ist, wie leicht zu sehen, $U_n = A_n$.

Die Gleichung (25) ist nun nichts anderes, als die Formel von *Lagrange*, die, wie wir früher in § 25 gesehen haben, eine ganze Funktion n^{ter} Ordnung linear durch $n+1$ vorgeschriebene Funktionswerte ausdrückt. Bezeichnet man also die besonderen Werte von t, welche den Ausdruck $f(a+th)$ in die einzelnen Größen der Reihe (24) verwandeln, mit $t_0, t_1, \cdots t_n$, so ist, wenn $f(x)$ als Funktion von t angesehen wird, nach § 25 (6)

$$U_n = [(t-t_0)(t-t_1)\cdots(t-t_{n-1})] : [(t_n-t_0)(t_n-t_1)\cdots(t_n-t_{n-1})]. \qquad (26)$$

Im vorliegenden Falle sind die Größen $t_0, t_1, \cdots$ identisch mit den Zahlen $0, 1, \cdots$, woraus sofort

$$A_n = U_n = [t(t-1)\cdots(t-n+1)] : n!$$

folgt, sodaß A_n mit dem Binomialkoeffizienten $(t)_n$ zusammenfällt. Nun sind aber, wie in § 26 zu den Koeffizienten A_q der allgemeinen Interpolationsformel J bemerkt worden ist, die A_q unabhängig von der Ordnung der dargestellten ganzen Funktion; folglich ist allgemein $A_q = (t)_q$. Damit nimmt die *Formel N* die nachstehende Gestalt (obere Zeichen) an:

$$\begin{aligned} f(a \pm th) = (a, 0) &\pm (t)_1 (a \pm \tfrac{1}{2}, 1) + (t)_2 (a \pm 1, 2) \\ &\pm (t)_3 (a \pm \tfrac{3}{2}, 3) + (t)_4 (a \pm 2, 4) \pm \cdots \end{aligned} \qquad (27)$$

Ersetzt man h durch $-h$, so kommt dies darauf hinaus, daß man das Differenzenschema um die Zeile a herum umklappt, wobei sich $(a+p, q)$ mit $(-1)^q (a-p, q)$ vertauscht. Durch Ausführung dieser Änderung gelangt man zu den unteren Zeichen in (27).

§ 32. Die für N benutzte Herleitung läßt sich auf die Formel G übertragen, welche die Gestalt

$$f(a+th) = A_0 (a, 0) + A_1 (a + \tfrac{1}{2}, 1) + A_2 (a, 2) + \cdots \qquad (28)$$

besitzt. Schreibt man die Funktionswerte, welche zur Bildung der vorkommenden Differenzen nötig sind, in derjenigen Reihenfolge hin, in welcher sie nach und nach zur Berechnung der in (28) aufeinanderfolgenden Differenzen herangezogen werden, so entsteht die Gliederfolge

$$(a, 0), \quad (a+1, 0), \quad (a-1, 0), \quad (a+2, 0), \quad (a-2, 0), \cdots, \qquad (29)$$

deren Bestandteile in der vorstehenden Anordnung wieder mit $y_0, y_1, \cdots$

bezeichnet werden sollen. Daraus ergeben sich die benutzten Differenzen in der Gestalt

$$\begin{aligned}
(a, 0) \quad &= y_0, \\
(a + \tfrac{1}{2}, 1) &= y_1 - y_0, \\
(a, 2) \quad &= y_1 - 2y_0 + y_2, \\
(a + \tfrac{1}{2}, 3) &= y_3 - 3y_1 + 3y_0 - y_2, \\
(a, 4) \quad &= y_3 - 4y_1 + 6y_0 - 4y_2 + y_4, \\
&\cdots\cdots\cdots\cdots\cdots\cdots
\end{aligned}$$

wo das Gesetz des Fortschreitens klar ist. Die Einsetzung dieser Ausdrücke führt wie vorhin auf die *Lagrange*sche Formel

$$f(a + th) = y_0 U_0 + y_1 U_1 + \cdots + y_n U_n,$$

in der wiederum U_n einerseits mit A_n zusammenfällt, andererseits durch den Quotienten (26) ausgedrückt ist, nur daß jetzt die Größen $t_0, t_1, \cdots$ wegen der in (29) vorgeschriebenen Reihenfolge der Funktionswerte $y_0, y_1, \cdots$ durch die Zahlen

$$0, \quad +1, \quad -1, \quad +2, \quad -2, \quad +3, \quad -3, \cdots$$

gegeben sind. Berechnet man nun den Quotienten der Reihe nach für die Fälle $n = 1, 2, 3, \cdots$, so entstehen die Binomialkoeffizienten

$$(t)_1, \quad (t)_2, \quad (t+1)_3, \quad (t+1)_4, \quad (t+2)_5, \quad (t+2)_6, \cdots.$$

Damit nimmt die *Formel G* die nachstehende Gestalt (obere Zeichen) an:

$$\begin{aligned}
f(a \pm th) = (a, 0) &\pm (t)_1 (a \pm \tfrac{1}{2}, 1) + (t)_2 (a, 2) \\
&\pm (t+1)_3 (a \pm \tfrac{1}{2}, 3) + (t+1)_4 (a, 4) \pm \cdots.
\end{aligned} \tag{30}$$

Die unteren Zeichen werden durch bloßes Umklappen des Differenzenschemas um die Zeile a erhalten.

Ersetzt man in der Formel G (untere Zeichen) t durch $-t$ und beachtet die Beziehung

$$(-t)_p = (-1)^p (t + p - 1)_p,$$

so entsteht die *Formel G′*, nämlich

$$\begin{aligned}
f(a + th) = (a, 0) &+ (t)_1 (a - \tfrac{1}{2}, 1) + (t+1)_2 (a, 2) \\
&+ (t+1)_3 (a - \tfrac{1}{2}, 3) + (t+2)_4 (a, 4) + \cdots.
\end{aligned} \tag{31}$$

Wird hierin auf der linken Seite a durch $a + h$ und t durch $t - 1$ ersetzt, so hat man auf der rechten Seite alle Zeilennummern um eine Einheit zu erhöhen und man erhält die *Formel G″*, nämlich

$$\begin{aligned}
f(a + th) = (a+1, 0) &+ (t-1)_1 (a + \tfrac{1}{2}, 1) + (t)_2 (a+1, 2) \\
&+ (t)_3 (a + \tfrac{1}{2}, 3) + (t+1)_4 (a+1, 4) + \cdots.
\end{aligned} \tag{32}$$

Nimmt man das Mittel aus G (obere Zeichen) und G', so entsteht die *Formel S*, nämlich

$$f(a+th)$$
$$= (a,0) + (t)_1(a,1) + (t+1)_3(a,3) + (t+2)_5(a,5) + \cdots \tag{33}$$
$$+ t\left[\tfrac{1}{2}(t)_1(a,2) + \tfrac{1}{4}(t+1)_3(a,4) + \tfrac{1}{6}(t+2)_5(a,6) + \cdots\right].$$

Hierin sind die Koeffizienten gerade oder ungerade Funktionen von t. Bildet man endlich noch das Mittel aus G (obere Zeichen) und G'', so gelangt man zu der *Formel B*, nämlich

$$f(a+th)$$
$$= (a+\tfrac{1}{2},0) + (t)_2(a+\tfrac{1}{2},2) + (t+1)_4(a+\tfrac{1}{2},4) + \cdots \tag{34}$$
$$+ (t-\tfrac{1}{2})\left[(a+\tfrac{1}{2},1) + \tfrac{1}{3}(t)_2(a+\tfrac{1}{2},3) + \tfrac{1}{5}(t+1)_4(a+\tfrac{1}{2},5) + \cdots\right].$$

Die Koeffizienten sind, wie man sich leicht überzeugt, gerade oder ungerade Funktionen von $t - \frac{1}{2}$. Bei der Rechnung werden die beiden niedrigsten Glieder gewöhnlich in der Gestalt

$$(a,0) + t(a+\tfrac{1}{2},1) \tag{35}$$

benutzt.

§ 33. Die Formeln N und G sind, wie sich gezeigt hat, bloße Umgestaltungen der Formel von *Lagrange*, und dasselbe gilt folgeweise auch von G' und G''. Die Interpolation nach diesen Formeln kommt also, wenn man jedesmal bis zur n^{ten} Differenz geht, darauf hinaus, daß man an die Stelle der vorgelegten Kurve $y = f(x)$ eine gewisse Parabel n^{ter} Ordnung setzt, die durch $n + 1$ in bestimmter Weise ausgewählte Punkte der gegebenen Kurve hindurchgeht. Sieht man nun zu, welche Ordinaten oder Funktionswerte bei den Formeln N, G, G', G'' benutzt werden, so gelangt man, unter Beiseitelassung der unteren Zeichen bei N und G, zu den nachstehenden vier Reihen:

$$\left.\begin{array}{llllll}
N\colon & (a,0) & (a+1,0) & (a+2,0) & (a+3,0) & (a+4,0)\cdots \\
G\colon & (a,0) & (a+1,0) & (a-1,0) & (a+2,0) & (a-2,0)\cdots \\
G'\colon & (a,0) & (a-1,0) & (a+1,0) & (a-2,0) & (a+2,0)\cdots \\
G''\colon & (a+1,0) & (a,0) & (a+2,0) & (a-1,0) & (a+3,0)\cdots
\end{array}\right\} \tag{36}$$

Nimmt man von diesen Reihen jedesmal die $n + 1$ niedrigsten Glieder, so erhält man die Ordinaten der Punkte, welche bei der Konstruktion der Parabeln für die Formeln N, G, G', G'' zu Grunde zu legen sind. Da ferner die Formel S als das Mittel von G und G' erhalten wurde, so entsteht die Parabel für S, wenn man aus den Ordinaten der beiden Parabeln G und G' das Mittel nimmt. Ebenso erhält man die Parabel für B, wenn man die Ordinaten der Parabeln G und G'' mittelt.

Beachtet man die Gruppierung der Funktionswerte in den obigen Reihen, so ergibt sich, daß bei G und G' die $n+1$ niedrigsten Glieder für ein gerades n auf dieselben Größen, nur in anderer Reihenfolge, führen. Demnach sind für ein gerades n die Parabeln G, G' und S identisch, woraus das Gleiche für die Formeln G, G' und S folgt. Ebenso fallen die Parabeln und die Formeln für G, G'' und B zusammen, sobald n ungerade ist.

Beschränkt man die Hilfsgröße t, die fortan als die „*Phase*" des Arguments $a+th$ bezeichnet werden soll, bei der Rechnung wie üblich auf echt gebrochene Werte, so besteht zwischen der Formel N und den übrigen Formeln ein ausgesprochener Unterschied bezüglich der Lage des eingeschalteten Punktes gegen die Punktreihe, die jedesmal zur Konstruktion der betreffenden Parabel dient. Denn bei N liegt der eingeschaltete Punkt nahe dem Ende der Punktreihe, bei den übrigen Formeln dagegen nahe der Mitte.

§ 34. Die rechnerische Benutzung der Formeln N, G, S, B kommt im wesentlichen darauf hinaus, daß man die Änderung ermittelt, welche die vorgelegte Funktion erfährt, wenn man von $f(a)$ auf das zu interpolierende $f(a+th)$ übergeht. Da es nun vorteilhaft ist Ziffern zu sparen, so wird man die zu berechnende Änderung oder mit anderen Worten die Größe th möglichst klein zu halten suchen. Demgemäß wählt man, wenn $f(x)$ gesucht wird, für a eines der beiden Tafelargumente, die den Wert x einschließen, und berechnet die Phase t aus

$$x = a + th \quad \text{oder} \quad t = (x - a) : h,$$

sodaß t stets als positiver oder negativer echter Bruch erscheint. Dieser Bruch läßt sich sogar zwischen den Grenzen ± 0.5 halten, wenn von den beiden für a in Betracht kommenden Werten derjenige gewählt wird, der dem Argument a am nächsten liegt.

Stehen die Werte von a und t fest, so hat man aus dem Differenzenschema die in der benutzten Formel auftretenden Differenzen $(a+p, q)$ zu entnehmen, weiter die zugehörigen, von t abhängenden Koeffizienten zu berechnen, endlich die Produkte aus Koeffizient und Differenz zu bilden und zu summieren.

Da dem Rechner vorläufig die Wahl zwischen den vier Formeln N, G, S, B freisteht, so gelangt man zu der Frage, welche Formel zu bevorzugen sei. Die naheliegende Antwort lautet, daß im allgemeinen die Formel mit den kleineren Koeffizienten die bequemere sein wird. Demnach handelt es sich jetzt darum, eine Vorstellung von dem Verhalten der Koeffizienten zu gewinnen. Zu dem Ende sind nachstehend die Koeffizienten zweiter bis sechster Ordnung tabuliert worden, und zwar für die runden Zehntel der Phase zwischen

den Grenzen -1 bis $+1$, unter Fortlassung der Phase $t = 0.0$, da für diese die betrachteten Koeffizienten sämtlich verschwinden. Die Überschrift gibt die Ordnungsnummer der Koeffizienten an, darunter folgen zeilenweise die Phase und die in Einheiten der vierten Decimale angesetzten Koeffizienten für die einzelnen Formeln N, G, S, B. Die Vorzeichen sind, weil es hier nur auf die absoluten Beträge ankommt, fortgelassen. Dafür sind den Buchstaben N, G, B die Zeichen $+$ oder $-$ angehängt, welche angeben, daß die Phase positiv oder negativ genommen worden ist. Bei S war eine solche Spaltung nicht nötig, weil die Koeffizienten gerade oder ungerade Funktionen der Phase sind.

2. Ordnung.

t	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
N_-	550	1200	1950	2800	3750	4800	5950	7200	8550	10000
N_+	450	800	1050	1200	1250	1200	1050	800	450	0
G_-	550	1200	1950	2800	3750	4800	5950	7200	8550	10000
G_+	450	800	1050	1200	1250	1200	1050	800	450	0
S	50	200	450	800	1250	1800	2450	3200	4050	5000
B_-	550	1200	1950	2800	3750	4800	5950	7200	8550	10000
B_+	450	800	1050	1200	1250	1200	1050	800	450	0

3. Ordnung.

t	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
N_-	385	880	1495	2240	3125	4160	5355	6720	8265	10000
N_+	285	480	595	640	625	560	455	320	165	0
G_-	165	320	455	560	625	640	595	480	285	0
G_+	165	320	455	560	625	640	595	480	285	0
S	165	320	455	560	625	640	595	480	285	0
B_-	110	280	520	840	1250	1760	2380	3120	3990	5000
B_+	60	80	70	40	0	40	70	80	60	0

4. Ordnung.

t	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
N_-	298	704	1233	1904	2734	3744	4953	6384	8058	10000
N_+	207	336	402	416	391	336	262	176	87	0
G_-	87	176	262	336	391	416	402	336	207	0
G_+	78	144	193	224	234	224	193	144	78	0
S	4	16	34	56	78	96	104	96	64	0
B_-	87	176	262	336	391	416	402	336	207	0
B_+	78	144	193	224	234	224	193	144	78	0

5. Ordnung.

t	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
N_-	245	591	1061	1676	2461	3444	4656	6129	7897	10000
N_+	161	255	297	300	273	228	173	113	54	0
G_-	33	63	89	108	117	116	104	81	45	0
G_+	33	63	89	108	117	116	104	81	45	0
S	33	63	89	108	117	116	104	81	45	0
B_-	10	25	42	60	78	92	96	87	58	0
B_+	6	9	8	4	0	4	8	9	6	0

6. Ordnung.

t	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
N_-	208	513	937	1508	2256	3215	4423	5924	7766	10000
N_+	132	204	233	230	205	168	124	79	37	0
G_-	17	34	49	61	68	70	64	51	30	0
G_+	16	30	40	47	49	47	40	30	16	0
S	1	2	4	7	10	12	12	11	7	0
B_-	17	34	49	61	68	70	64	51	30	0
B_+	16	30	40	47	49	47	40	30	16	0

An der Hand der vorstehenden Zahlen lassen sich jetzt bestimmte Aussagen über die Zweckmäßigkeit der einzelnen Formeln machen.

§ 35. Vergleicht man zunächst innerhalb derselben Formel und Ordnung die Koeffizienten, die zu entgegengesetzten Phasenwerten gehören, so findet natürlich Übereinstimmung in denjenigen Fällen statt, in denen der Koeffizient, wie das z. B. bei S durchgehends der Fall ist, sich als gerade oder ungerade Funktion der Phase darstellt. In allen anderen Fällen stellt sich dagegen die negative Phase ungünstiger als die positive. Man wird daher, von der Formel S abgesehen, die Phase positiv wählen oder aber, wenn die Phase negativ ist, bei N und G mit den unteren Zeichen der beiden in (27) und (30) angesetzten Formeln rechnen.

Beschränkt man sich jetzt bei der weiteren Vergleichung auf positive Phasenwerte und hält für jede Ordnung die Zahlen N, G, B gegeneinander, so gelangt man zu den Urteilen: „B teils gleich G, teils besser als G“ und „G teils gleich N, teils ausgesprochen besser als N“. Hiernach ist von den drei Formeln N, G, B im allgemeinen B zu bevorzugen.

Bei S ist es offenbar vorteilhaft, die Phase nicht zwischen 0 und 1, sondern zwischen -0.5 und $+0.5$ anzunehmen. Vergleicht man unter dieser Voraussetzung S mit B, so ergibt sich, daß S abwechselnd

besser oder schlechter als B ausfällt, je nachdem die Ordnungsnummer gerade oder ungerade ist. Dem entsprechend wird sich der Einfluß der niedrigsten in der Rechnung vernachlässigten Differenz bei S kleiner (größer) als bei B stellen, wenn die Ordnung dieser Differenz gerade (ungerade) ist. Infolgedessen wird man, je nach den Umständen, S oder B zu bevorzugen haben.

Betrachtet man die geraden und ungeraden Ordnungen getrennt für sich, so zeigen die Beträge der Koeffizienten eine ausgesprochene Abnahme bei steigender Ordnungsnummer. Infolgedessen ist es gewöhnlich nicht nötig, die letzte noch merkliche Differenz des Differenzenschemas bei der Interpolation wirklich zu berücksichtigen. Soll z. B. aus dem in § 15 als Tabelle IV gegebenen Täfelchen interpoliert werden, so hat man bei Benutzung von S nur die dritten Differenzen mitzunehmen; ferner würde man sich bei Benutzung von B, falls auf die Einheit der achten Stelle kein besonderes Gewicht gelegt wird, mit den zweiten Differenzen begnügen dürfen.

§ 36. Da die Berechnung der Koeffizienten, namentlich bei den höheren Differenzen, gewöhnlich den lästigsten Teil der Interpolationsarbeit bildet, so sind die Werte der Koeffizienten vielfach tabuliert worden. Solche Tafeln, besonders für die Formeln S und B, finden sich in den meisten Tafelsammlungen, die für Zwecke der Astronomie oder Geodäsie zusammengestellt worden sind, so u. a. in

Formeln und Hülfstafeln für geographische Ortsbestimmungen von Th. Albrecht. (1894 in dritter Auflage erschienen);
Tafeln zur theoretischen Astronomie bearbeitet von J. Bauschinger.

Hat man keine Hilfstafeln zur Hand, so ist es im allgemeinen zweckmäßig, nach der Formel G zu rechnen, trotz der im allgemeinen etwas langsameren Abnahme der Koeffizienten. Beachtet man nämlich den Umstand, daß bei G jeder Koeffizient Teiler aller folgenden Koeffizienten ist, so kann man die beiden in (30) angesetzten Formeln in folgender Weise aufbauen. Man bildet zunächst die Größen

$$\begin{aligned} A = t, \quad B &= (t-1):2, \quad C = (t+1):3, \\ D &= (t-2):4, \quad E = (t+2):5, \end{aligned} \tag{37}$$

u. s. w.

und rechnet dann, je nach dem Vorzeichen der Phase, entweder nach den Gleichungen

$$\begin{aligned} f(a+th) = (a, 0) + A \cdot \mathrm{I}, \qquad & \mathrm{I} = (a+\tfrac{1}{2}, 1) + B \cdot \mathrm{II}, \\ \mathrm{II} = (a, 2) + C \cdot \mathrm{III}, \quad & \mathrm{III} = (a+\tfrac{1}{2}, 3) + D \cdot \mathrm{IV}, \\ \mathrm{IV} = (a, 4) + E \cdot \mathrm{V}, \quad & \mathrm{V} = (a+\tfrac{1}{2}, 5) + F \cdot \mathrm{VI}, \end{aligned} \tag{38}$$

u. s. w.

oder aber nach

$$f(a-th) = (a,0) - A\cdot \mathrm{I}, \qquad \mathrm{I} = (a-\tfrac{1}{2}, 1) - B\cdot \mathrm{II},$$
$$\mathrm{II} = (a, 2) - C\cdot \mathrm{III}, \quad \mathrm{III} = (a-\tfrac{1}{2}, 3) - D\cdot \mathrm{IV}, \qquad (39)$$
$$\mathrm{IV} = (a, 4) - E\cdot \mathrm{V}, \qquad \mathrm{V} = (a-\tfrac{1}{2}, 5) - F\cdot \mathrm{VI},$$

u. s. w.

Die Berechtigung zu diesen beiden Ansätzen erhellt sofort, wenn man die Hilfsgrößen I, II, ·· wieder eliminiert.

Denkt man sich bei der Benutzung von (38) in dem Differenzenschema zwischen den beiden zur Verwendung kommenden Zeilen, nämlich der Hauptzeile a und der Zwischenzeile $a + \frac{1}{2}$ einen Strich gezogen, so geht aus jeder Spalte diejenige Differenz in die Rechnung ein, welche dem Striche am nächsten anliegt. Faßt man ferner die Größen I, II, ·· als „verbesserte" Differenzen und die Produkte $A\cdot \mathrm{I}$, $B\cdot \mathrm{II}$, ·· als die angebrachten „Verbesserungen" auf, so wird, wie die Betrachtung des Vorzeichens in den einzelnen Gleichungen erkennen läßt, im allgemeinen jede Differenz durch die angebrachte Verbesserung der ihr an der anderen Seite des Striches gegenüberstehenden Differenz genähert. Dieselbe Bemerkung gilt auch für den Ansatz (39), nur daß hier der Strich zwischen den Zeilen a und $a - \frac{1}{2}$ zu ziehen ist. Diese *Strichregel* wird gelegentlich empfohlen, und zwar mit der Begründung, daß man bei ihrer Befolgung von der Berücksichtigung der Vorzeichen absehen kann; sie ist in der Tat ganz nützlich, jedoch ist zu beachten, daß sie versagt, wenn die beiden einander gegenüberstehenden Differenzen denselben Wert besitzen, ohne daß die anzubringende Verbesserung null ist.

Eine andere Regel, die bei der Formel G empfohlen wird, beruht darauf, daß die in (37) zusammengestellten Faktoren beim weiteren Fortschreiten gegen den Wert $\frac{1}{2}$ konvergieren. Die Verbesserung der q^{ten} Differenz ist also für ein größeres q nahe gleich der Hälfte der $(q+1)^{\text{sten}}$ Differenz. Da nun die Verbesserung nach dem Striche hin erfolgt, so wird, falls die Formel nur bis zur q^{ten} Differenz benutzt werden soll, die $(q+1)^{\text{ste}}$ Differenz genähert mit berücksichtigt, wenn man als q^{te} Differenz statt des dem Striche zunächst liegenden Wertes das Mittel aus den beiden dem Striche anliegenden Werten benutzt.

§ 37. Die Formel G war in doppelter Gestalt angesetzt worden, sodaß die Berechnung von $f(a+th)$ auf zwei verschiedene Arten erfolgen kann; dieser Umstand läßt sich als nützliche Kontrole verwenden. Soll z. B. aus der Tabelle IV in § 15 der Logarithmus von 52.34567 interpoliert werden, so hat man die beiden Ansätze

obere Zeichen: $a = 52 \qquad t = 0.34567$
untere Zeichen: $a = 53 \qquad t = 0.65433$

zur Verfügung. Die Rechnung liefert dann, wenn man die Differenzen der Ordnungen 5, 6, · · vernachlässigt und mit dem Mittel der vierten Differenzen beginnt, die nachstehenden Zahlen, die bei den Differenzen in Einheiten der siebenten Decimale angesetzt und im übrigen ihrer Bedeutung nach durch die linker Hand stehenden Angaben gekennzeichnet sind.

A	+ 0.34567	+ 0.65433
B	− 0.3272	− 0.1728
C	+ 0.449	+ 0.551
D	− 0.41	− 0.34
IV	− 4.5	− 4.5
D·IV	+ 1.8	+ 1.5
$(a \pm \frac{1}{2}, 3)$	+ 58.0	+ 58.0
III	+ 59.8	+ 56.5
C·III	+ 26.9	+ 31.1
$(a, 2)$	− 1605.0	− 1547.0
II	− 1578.1	− 1578.1
B·II	+ 516.3	+ 272.7
$(a \pm \frac{1}{2}, 1)$	+ 82726.0	+ 82726.0
I	+ 83242.3	+ 82453.3
A·I	+ 28774.4	+ 53951.7
$(a, 0)$	1.7160033.0	1.7242759.0
$f(a \pm th)$	1.7188807.4	1.7188807.3

Der „Thesaurus“ von *Vega* gibt 1.7188807.6, sodaß beim Abrunden der interpolierten Werte ein Fehler von 0.6 Einheiten der siebenten Stelle entstehen würde.

§ 38. Die Formeln G, S, B versagen, wenn die Interpolation in einem Endintervall der vorgelegten Tafel auszuführen ist. Soll z. B. aus dem in Tabelle IV enthaltenen Täfelchen der Logarithmus 50.1 gefunden werden, so fehlen schon die für die genannten Formeln nötigen Differenzen zweiter Ordnung. Häufig hilft man sich nun in solchen Fällen dadurch, daß man die fehlenden Differenzen durch eine sogenannte „*Extrapolation*“ ergänzt. In dem nachstehenden Täfelchen sind zunächst ohne Klammern die Zahlen angesetzt, die man für die Aufsuchung von Log 50.1 unmittelbar aus Tabelle IV entnehmen kann.

x	Log	I	II	III	IV
50	1.6989700		(− 1741)		(− 4)
		+ 86002		(+ 70)	
51	1.7075702		− 1671		(− 4)
		+ 84331		+ 66	
52	1.7160033		− 1605		− 8

Nimmt man nun an, daß für die in Betracht kommende Gegend der Tafel die fünften Differenzen unmerklich, also die vierten Differenzen merklich konstant seien, so kann man die vorhandenen vierten Differenzen der Tabelle IV benutzen, um diesen konstanten Wert abzuleiten. Bildet man aus den drei vorhandenen Differenzen, nämlich -8, -1, -3 das arithmetische Mittel, um die von den Abrundungsfehlern herrührenden Sprünge etwas auszugleichen, so erhält man den Wert -4, der oben, eingeklammert, an dem Orte der beiden fehlenden vierten Differenzen eingetragen worden ist. Daraus ergeben sich dann sofort die ebenfalls eingeklammerten Zahlen für die anderen noch fehlenden Differenzen, und man kann nun ohne weiteres nach der Formel G oder B rechnen.

Es versteht sich von selbst, daß eine solche Extrapolation stets mit einer gewissen Vorsicht anzuwenden ist. Deshalb wird man, wenn die Extrapolation aus irgend welchen Gründen als bedenklich erscheint, bei den Endintervallen des Differenzenschemas lieber seine Zuflucht zu der in (27) angesetzten Formel N nehmen, die für den Anfang oder den Schluß des Schemas brauchbar ist, je nachdem man die oberen oder die unteren Zeichen benutzt. Sind für die dort auftretenden Binomialkoeffizienten keine Hilfstafeln zur Hand, so legt man die Rechnung ähnlich wie bei G an. Man bildet zunächst die Faktoren

$$A = t, \qquad B = (t-1):2, \qquad C = (t-2):3, \cdots \tag{40}$$

und rechnet dann nach den Gleichungen

$$\left.\begin{aligned} f(a \pm th) &= (a, 0) \pm A\cdot\mathrm{I}, \\ \mathrm{I} &= (a \pm \tfrac{1}{2}, 1) \pm B\cdot\mathrm{II}, \\ \mathrm{II} &= (a \pm 1, 2) \pm C\cdot\mathrm{III}, \\ &\text{u. s. w.} \end{aligned}\right\} \tag{41}$$

wo die oberen (unteren) Zeichen für den Anfang (Schluß) des Differenzenschemas anzuwenden sind.

Ein anderer Fall, bei dem die Anwendung der Formel N in Betracht kommt, ist folgender. Wenn wenige Intervalle vor dem einzuschaltenden Werte $f(a + th)$ eine irreguläre Stelle liegt, so versagen die Formeln G, S, B, weil sich die Wirkung der Irregularität in den Differenzen der Zeile a oder der anliegenden Zwischenzeile schon bei einer niedrigen Ordnungsnummer geltend macht. Dagegen sind die Glieder der absteigenden Treppe, nämlich

$$(a, 0), \qquad (a + \tfrac{1}{2}, 1), \qquad (a + 1, 2), \qquad \text{u. s. w.,}$$

der Wirkung jener Irregularität nicht unterworfen, sodaß die Formel N anwendbar bleibt.

§ 39. Die entwickelten Formeln arbeiten völlig sicher, sobald für die vernachlässigte Korrektion K die früher besprochenen Voraussetzungen der Unmerklichkeit erfüllt sind. Man könnte deshalb z. B. eine siebenstellige Tafel für Log x auf die 900 Werte für die ganzzahligen x von 100 bis 1000 beschränken und würde bei der Formel B gleichwohl nur die zweiten Differenzen zu berücksichtigen haben, da die dritten Differenzen einer solchen Tafel unter 10 Einheiten der siebenten Stelle bleiben und deshalb nach Ausweis der in § 34 zusammengestellten Koeffizientenwerte bedeutungslos werden. Trotzdem hat man bei den vorhandenen Tafeln ein hundertmal kleineres Intervall gewählt, denn beim Aufsuchen des Logarithmus oder Numerus fällt der größere Teil des Zeitaufwandes auf die Interpolation; darum ist es wesentlich, die Interpolationsarbeit möglichst abzukürzen. Selbst wenn man nur die ersten Differenzen zu berücksichtigen hat, machen sich die Unterschiede der Intervallgröße oder, was auf dasselbe hinauskommt, die Unterschiede in den Beträgen der ersten Differenzen recht fühlbar. Aus diesem Grunde wäre es eigentlich von jeher zweckmäßiger gewesen, nicht den Logarithmus, sondern die inverse Funktion „Numerus $x = 10^x$“ zu tabulieren. Bei der gewöhnlichen siebenstelligen Einrichtung für Log x gehen nämlich die ersten Differenzen von 43 bis 435, und die Menge der dreiziffrigen Differenzen beträgt rund 48 %. Bei der entsprechenden Tabulierung des Numerus würden dagegen die Differenzen von 22 bis 231 gehen, und die Menge der dreiziffrigen Differenzen betrüge noch nicht 37 %, während der Umfang der Tafel von 180 Seiten nur auf 200 Seiten stiege.

Selbstverständlich findet die zur Erleichterung der Interpolation vorgenommene Verkleinerung des Intervalls eine Grenze an dem räumlichen Umfange der Tafel und an den Kosten der Herstellung. Deshalb trifft man z. B. in den Ephemeridensammlungen Tafeln, bei denen, weil es sich um seltener gebrauchte Größen handelt, mit vierten oder gar fünften Differenzen zu interpolieren ist.

Die Gründe, welche bei der Benutzung einer fertigen Tafel für die möglichste Vermeidung einer mehrgliedrigen Interpolation sprechen, verkehren sich in ihr Gegenteil, wenn es sich nicht um die *Benutzung*, sondern um die *Herstellung* einer Tafel handelt. Wenn z. B. bei der mit dem Intervall h ausgeführten Tabulierung von $f(x)$ für den einzelnen Funktionswert durchschnittlich der Zeitaufwand X erforderlich ist, so verlangt die Tafelstrecke $H = Nh$ den Aufwand NX. Legt man dagegen die Tafel zunächst mit dem „groben“ Intervall H an und bringt das Intervall nachher durch $N - 1$ Interpolationen auf das „feine“ Intervall h, so ergibt sich, wenn die einzelne Interpolation durchschnittlich den Aufwand Y erfordert, für die Strecke H der Aufwand $X + (N - 1)Y$. Der Unterschied ist also $(N - 1)(X - Y)$, und zwar zu Gunsten der Interpolation, sobald X größer als Y ist.

Dieser Fall ist aber die Regel. Die Differenz $X - Y$ ist häufig so bedeutend, daß es vorteilhaft ist, für N ziemlich große Werte, z. B. 200 oder 300, zu wählen, obgleich bei wachsendem N die Anzahl der in der Interpolationsformel mitzunehmenden Glieder wächst, so daß Y zunimmt und $X - Y$ abnimmt.

§ 40. Soll der Übergang von dem ursprünglichen groben Intervall $H = Nh$ zu dem Intervall h auf dem gewöhnlichen Wege ohne Maschine erfolgen, so wird dazu gern die sogenannte *„Interpolation in die Mitte"* benutzt. Die Formel hierfür entsteht, wenn man in der Formel B für die Phase t den Wert $\frac{1}{2}$ setzt. Beachtet man den Umstand, daß die Koeffizienten von B gerade oder ungerade Funktionen von $t - \frac{1}{2}$ sind, so erhält man

$$f(a + \tfrac{1}{2}h) = (a + \tfrac{1}{2}, 0) - \tfrac{1}{8}(a + \tfrac{1}{2}, 2) + \tfrac{3}{128}(a + \tfrac{1}{2}, 4) - \tfrac{5}{1024}(a + \tfrac{1}{2}, 6) + \cdots \tag{42}$$

Die vorstehende Formel besitzt offenbar den Vorzug, daß die ungeraden Differenzen herausfallen, und daß die Koeffizienten bequeme Werte besitzen. Dem entsprechend wählt man, nachdem das endgültige Intervall h festgesetzt worden ist, für N eine passende Potenz von 2 und rechnet die Tafel zunächst für das grobe Intervall $Nh = H$. Dann bringt man durch wiederholte Anwendung von (42) das Intervall der Reihe nach auf die Beträge $H : 2$, $H : 4$, u. s. w., bis man bei dem Intervall h angelangt ist.

Schreibt man (42) für den Augenblick in der Gestalt

$$f(a + \tfrac{1}{2}h) = (a + \tfrac{1}{2}, 0) + C,$$

so treten offenbar nach ausgeführter Halbierung des Intervalls an die Stelle der ursprünglichen Differenz erster Ordnung die beiden neuen Differenzen erster Ordnung

$$\begin{aligned} f(a + \tfrac{1}{2}h) - f(a) &= \tfrac{1}{2}(a + \tfrac{1}{2}, 1) + C, \\ f(a + h) - f(a + \tfrac{1}{2}h) &= \tfrac{1}{2}(a + \tfrac{1}{2}, 1) - C. \end{aligned} \tag{43}$$

Die Rechnung kommt also für die Differenzen erster Ordnung darauf hinaus, daß man die ursprüngliche Differenz $(a + \frac{1}{2}, 1)$ in die beiden vorstehend angegebenen Bestandteile spaltet. Da nun mit den Differenzen erster Ordnung zugleich die der höheren Ordnungen gegeben sind, so kann man, ohne erst die Funktionswerte selber zu bilden, die nächste Halbierung des Intervalls durch bloßes Spalten der Größen (43) bewerkstelligen u. s. w., bis man bei der letzten Halbierung angelangt ist. Ist diese Rechnung erledigt, so ergeben sich die Funktionswerte durch bloßes Aufsummieren der gefundenen Differenzen erster Ordnung; man hat nur darauf zu achten, daß bei der Spaltung jedesmal die Summe der beiden neuen Differenzen genau gleich dem Werte der alten ist.

§ 41. Kommt es bei der Tabulierung einer Funktion auf die Korrektheit der letzten Stelle an, so hat man die ursprüngliche grobe Tafel mit erheblich mehr Decimalen anzulegen, als schließlich bei behalten werden; so sind z. B. bei den siebenstelligen logarithmischen Tafeln von den Berechnern gewöhnlich sechs bis acht Decimalen zugegeben worden. Infolgedessen treten vielziffrige Zahlen auf, für deren Bewältigung gerade die Rechenmaschine am Platze ist. Da hierbei die fortgesetzte Halbierung des Intervalls keine besonderen Vorteile bieten würde, zumal man nach jeder Halbierung die höheren Differenzen von neuem zu bilden hat, so ist es zweckmäßiger, den Übergang von dem groben auf das endgültige feine Intervall in einem Zuge zu bewerkstelligen. Selbstverständlich hat sich dabei die Anlage der Rechnung sowohl nach der Eigenart der Maschine, als auch nach dem Verhalten der Differenzen in der groben Tafel zu richten, sodaß sich nicht wohl feste für alle Fälle gültige Regeln aufstellen lassen. Deshalb sollen hier nur zwei Beispiele behandelt werden, und zwar unter der Voraussetzung, daß die Maschine nach Art des *Thomas*-Modells arbeitet.

Die nachstehend in Einheiten der siebenten Stelle angesetzten Zahlen gehören zwei benachbarten Hauptzeilen und der von ihnen eingeschlossenen Zwischenzeile einer vorgelegten Tafel an, die auf zehnmal kleineres Intervall gebracht werden soll; die Überschriften geben die Ordnungsnummern an.

(0)	(1)	(2)	(3)
+ 2442958		− 473	
	+ 107267		− 19
+ 2550225		− 492	

Ferner soll die neue Tafel nur fünfstellig angesetzt werden, mit dem Hinzufügen, daß die normale obere Grenze der Abrundung, nämlich 5 Einheiten der sechsten Stelle, hier und da um eine Kleinigkeit überschritten werden dürfe. Zieht man nun die in § 34 gegebene Zusammenstellung der Koeffizientenwerte zu Rate, so ergibt sich, daß der Einfluß der dritten Differenz bei der Formel S die Einheit der siebenten Stelle überschreitet, bei der Formel B dagegen unter 2 Einheiten der achten Stelle bleibt. Demgemäß kann man bei Benutzung von B die dritten Differenzen unbedenklich vernachlässigen, also gemäß (34) und (35) die Verkleinerung des Intervalls nach der Formel

$$f(a+th) = (a, 0) + t(a+\tfrac{1}{2}, 1) + \frac{t(t-1)}{1\cdot 2}(a+\tfrac{1}{2}, 2) \qquad (44)$$

bewirken. Denkt man sich hierin für t die 11 Werte von 0.0 bis 1.0 gesetzt, so entsteht ein Täfelchen, das außer den bereits vorhandenen Funktionswerten 9 dazwischen geschaltete enthält. Bildet man zu

diesem Täfelchen die ersten und zweiten Differenzen, so ist das erste und letzte Glied in der Spalte der ersten Differenzen durch die Ausdrücke

$$\begin{aligned} &0.1\,(a+\tfrac{1}{2},\,1) - 0.045\,(a+\tfrac{1}{2},\,2) \\ &0.1\,(a+\tfrac{1}{2},\,1) + 0.045\,(a+\tfrac{1}{2},\,2) \end{aligned} \tag{44}$$

gegeben, wogegen die zweiten Differenzen den konstanten Wert

$$0.01\,(a+\tfrac{1}{2},\,2) \tag{45}$$

besitzen. Hiermit nimmt die Rechnung die aus nachstehender Tabelle ersichtliche Gestalt an, wobei die einzelnen Zahlen in Einheiten der elften Decimale angesetzt sind, um durch die Interpolation keine Abrundung in die Rechnung hineinzubringen. Zunächst sind in das durch Einschreiben der Phasenwerte und der Ordnungsnummern vorbereitete Schema in der Spalte (0) an erster und letzter Stelle die beiden bereits bekannten und zu den Phasen 0.0 und 1.0 gehörigen Funktionswerte eingetragen, und zwar mit einer nachher zu erläuternden Erhöhung um 5 Einheiten der sechsten Stelle. Weiter sind in Spalte (1) am Anfang und am Ende die beiden aus (44) folgenden Werte für die Differenzen erster Ordnung eingesetzt. Endlich ist unter (2) die konstante Differenz zweiter Ordnung nach (45) eingetragen.

t	(0)	(1)	(2)
0.0	+ 244 3008 0000		
		+ 107484 125	
0.1	+ 245 3756 4125		— 48 250
		+ 107435 875	
0.2	+ 246 4500 0000		
		+ 107387 625	
0.3	+ 247 5238 7625		
		+ 107339 375	
0.4	+ 248 5972 7000		
		+ 107291 125	
0.5	+ 249 6701 8125		
		+ 107242 875	
0.6	+ 250 7426 1000		
		+ 107194 625	
0.7	+ 251 8145 5625		
		+ 107146 375	
0.8	+ 252 8860 2000		
		+ 107098 125	
0.9	+ 253 9570 0125		
		+ 107049 875	
1.0	+ 255 0275 0000		

Nach den beschriebenen Vorbereitungen wird nun zunächst die Spalte (1) ausgefüllt, indem man zu dem Anfangsgliede von (1) wiederholt den Wert von (2) hinzufügt. Dieses Aufsummieren muß

am Schlusse der Spalte auf den daselbst bereits eingetragenen Wert führen. Dann wird in der gleichen Weise die Spalte (0) durch Aufsummieren der ersten Differenzen ausgefüllt, wobei in der letzten Zeile wiederum der bereits eingetragene Wert zum Vorschein kommen muß.

Das beschriebene Verfahren wickelt sich, wie eine mehrfache Erprobung bei der Verfeinerung ausgedehnter Tafeln gelehrt hat, äußerst glatt ab, sobald zwei Personen zusammenarbeiten, eine an der Maschine, die andere mit der Feder. Die zur Vermeidung der Abrundung mitgenommenen Decimalen, die sich bei der gewöhnlichen Handarbeit als ein äußerst lästiger Ziffernballast geltend machen würden, verlieren diese Eigenschaft, sobald mit einer geeigneten Maschine gearbeitet wird. Wichtiger aber sind noch die durchgreifenden Kontrollen und der Umstand, daß diese Kontrollen absolut stimmen müssen.

Der Zuschlag von 5 Einheiten der sechsten Stelle, der in dem Beispiel bei den Funktionswerten vorgenommen worden ist, bewirkt offenbar, daß man sich beim Niederschreiben der verlangten fünfstelligen Funktionswerte um die von der Maschine angezeigten 6 letzten Stellen nicht weiter zu kümmern braucht, denn die fünfte Stelle besitzt bereits den Wert, den sie ohne jenen Zuschlag bei richtiger Abrundung anzunehmen hat. Dieser sehr nützliche Kunstgriff ist meines Wissens zuerst von *J. F. von Wrede* angegeben worden.

§ 42. Hat man höhere Differenzen zu berücksichtigen, so wird das beschriebene Summenverfahren gewöhnlich unausführbar, weil die Maschine nicht mehr die nötige Stellenzahl hergibt. Man hat dann die Rechnung mit der Maschine anders anzulegen.

Die hier in Betracht kommenden Interpolationsformeln besitzen die Gestalt

$$f(a + th) = D_0 + A_1 D_1 + A_2 D_2 + \cdots, \tag{46}$$

worin die Differenzen D unmittelbar dem Differenzenschema der vorgelegten Tafel zu entnehmen sind, während die von der Phase t abhängenden Faktoren A ein ebenfalls bekanntes Größensystem bilden, sobald man festgesetzt hat, wie weit die groben Intervalle der vorgelegten Tafel verfeinert werden sollen. Ferner bleiben bei der Durchrechnung von (46) die Differenzen D innerhalb eines groben Intervalls konstant, wogegen die Faktoren A sich von einem groben Intervall zum anderen wiederholen.

Bei Anwendung der Maschine könnte man zunächst nun so verfahren, daß man nach (46) jedes $f(a + th)$ für sich in einem Zuge fertig rechnet. Hierbei würde man jedoch, ganz abgesehen von dem in § 6 erwähnten Übelstande, den Vorteil aus der Hand geben, der

darin liegt, daß in der Gesamtheit der zu bildenden Produkte AD beide Faktoren wiederholt vorkommen. Aus diesem Grunde ist es vorzuziehen, durch verkettete Multiplikation in einem Zuge zunächst immer diejenigen Produkte zu bilden, welche z. B. in dem Faktor A übereinstimmen. Sind diese Produktgruppen fertig gerechnet und in das Rechenschema eingetragen, so erfolgt die schon wegen etwa begangener Schreibfehler unerläßliche Kontrolle dadurch, daß man die sämtlichen Produkte nochmals verkettet durchrechnet, nun aber in Gruppen, innerhalb deren der Faktor D konstant ist. Hierbei fällt das Niederschreiben der in der Maschine erscheinenden Zahlen fort, da ja dieselben bereits in das Rechenschema eingetragen sind. Ist das System der Produkte fertig gestellt, so folgt deren Summierung nach (46), dann die Abrundung der Summen, endlich die Schlußprüfung durch Differenzen.

§ 43. Die Abrundungsfehler der Funktionswerte machen sich, wie wir in § 14 gesehen haben, in dem Differenzenschema um so stärker geltend, je höher die Ordnungsnummer einer Differenz ist. Wenn nun auch diese Fehlervergrößerung beim Interpolieren durch das Verhalten der Koeffizienten wieder vermindert wird, so läßt sich doch ohne nähere Untersuchung nicht übersehen, wie die Fehler in dem Werte von $f(a + th)$ wirken, und ob nicht etwa gerade bei der Benutzung höherer Differenzen merkliche Fehlerreste übrig bleiben können. Stellt man demgemäß die Aufgabe, der Fehlerwirkung genauer nachzugehen, so kann man einmal nach der Ungenauigkeit fragen, die „im Durchschnitt“ zu befürchten ist, sodann aber auch nach dem Betrage, den die Wirkung im ungünstigsten Falle erreicht. Die erste Frage soll hier beiseite bleiben, weil sie umständlicher zu beantworten ist und für den Rechner ohnehin ein geringeres Interesse besitzt als die zweite.

Bei der Untersuchung gehen wir von den Formeln N, G, G', G'' aus, die bis zur n^{ten} Differenz angesetzt sein mögen; außerdem sollen in den Formeln N und G, was hier genügt, nur die oberen Zeichen der Darstellungen (27) und (30) Verwendung finden. Um die Parabeln n^{ter} Ordnung, die zu den genannten Formeln gehören, zu erhalten, hat man als Ordinaten aus dem Differenzenschema von $f(x)$ die $n + 1$ aufeinanderfolgenden Funktionswerte $(a + p, 0)$ zu entnehmen, die jedesmal durch die $n + 1$ niedrigsten Glieder der in (36) zusammengestellten Reihen gegeben sind. Ist in diesen Ordinatenreihen $a + r$ die niedrigste Zeilennummer, $a + s$ die höchste, so wird $s - r = n$. Ferner wird, da n mindestens gleich Eins ist, nach Ausweis der Zusammenstellung (36) r niemals positiv, s niemals negativ. Setzt man also fest, daß die Nummer p im folgenden stets von r bis s läuft, so ist in dieser Nummernreihe jedesmal auch der Wert $p = 0$ enthalten.

Um die Ausdrücke für N, G, G', G'' in der *Lagrange*schen Form zu erhalten, hat man das Produkt

$$T(t) = (t - r)(t - r - 1) \cdots (t - s + 1)(t - s)$$

nebst seiner Ableitung $T'(t)$ aufzustellen, dann die Ausdrücke

$$V(t, p) = T(t) : T'(p), \qquad U(t, p) = V(t, p) : (t - p) \tag{47}$$

anzusetzen und endlich den Ausdruck

$$f(a + th) = \sum_p (a + p, 0)\, U(t, p) \tag{48}$$

zu bilden. Ersetzt man hierin die Funktionswerte $(a + p, 0)$ durch ihre Fehler, so erhält man statt $f(a + th)$ den Fehler dieser Größe. Oder mit anderen Worten: der Fehler der Interpolation wird gefunden, wenn man das Differenzenschema der Fehler bildet und darin interpoliert.

Von den Fehlern der Funktionswerte $(a + p, 0)$ soll jetzt vorausgesetzt werden, daß sie zwischen den Grenzen $\pm z$ enthalten seien. Dann wird der Maximalfehler der Interpolation, der mit Pz bezeichnet werden soll, aus (48) erhalten, wenn man darin die Größen $(a + p, 0)$ durch $\pm z$ ersetzt und die Vorzeichen von z derart wählt, daß die Glieder der Summe (48) durchweg das gleiche Vorzeichen annehmen. Ersetzt man in derselben Weise die Größen $(a + p, 0)$ nicht durch $\pm z$, sondern durch ± 1, so liefert die Summe (48) die Zahl P selber, die man als die „extreme Fehlervergrößerung" bezeichnen kann.

Die letzte Bemerkung soll noch etwas anders gefaßt werden. Bezeichnet das Symbol $\mathrm{sg}(x)$ die Zahl $+1$ oder 0 oder -1, je nachdem x positiv oder null oder negativ ist, so kann man sagen: Die Summe (48) liefert die Zahl P, wenn man die Größen $(a + p, 0)$ durch $\mathrm{sg}[U(t, p)]$ ersetzt. Die Lösung der Aufgabe kommt also jetzt darauf hinaus, daß man aus den Größen $\mathrm{sg}[U(t, p)]$ das Differenzenschema bildet und darin für das Argument $a + th$ interpoliert. Demnach sind zunächst die sg-Größen aufzustellen.

§ 43. Das Produkt $T(t)$ geht nur an den Stellen $t = p$ durch Null hindurch, und zwar abwechselnd von der positiven zur negativen Seite, und umgekehrt. Demnach weist die Reihe der Größen $T'(p)$ lauter Zeichenwechsel auf. Dasselbe gilt von der Reihe der in (47) aufgestellten Größen $V(t, p)$ bei festgehaltenem t. Beschränkt man, wie üblich, in dem zur Interpolation benutzten Argument $a + th$ die Phase t auf positive echt gebrochene Werte und bezeichnet mit u irgend einen andern hiernach zulässigen Phasenwert, so ist der Quotient $uT(t) : T(u)$ positiv und behält diese Eigenschaft auch noch bei, wenn u gegen Null geht. Für $u = 0$ nimmt aber der genannte Quotient den Wert $V(t, 0)$ an, sodaß $V(t, 0)$ sicher positiv ist.

Damit sind sofort auch die Vorzeichen der übrigen $V(t, p)$ bestimmt.

Das Vorzeichen von $t - p$ ist positiv oder negativ, je nachdem p der Zahlenreihe $0, -1, -2, \cdots$ oder $1, 2, 3, \cdots$ angehört. Beachtet man nun noch, daß wegen (47)

$$\operatorname{sg}[U(t, p)] = \operatorname{sg}[V(t, p)] \cdot \operatorname{sg}(t - p)$$

ist, so kann man sofort folgende Tabelle bilden:

$$\begin{array}{rcccccccc}
p = & \cdots -3 & -2 & -1 & 0 & +1 & +2 & +3 \cdots \\
\operatorname{sg}(V) = & \cdots -1 & +1 & -1 & +1 & -1 & +1 & -1 \cdots \\
\operatorname{sg}(t-p) = & \cdots +1 & +1 & +1 & +1 & -1 & -1 & -1 \cdots \\
\operatorname{sg}(U) = & \cdots -1 & +1 & -1 & +1 & +1 & -1 & +1 \cdots
\end{array}$$

Die ungünstigste Fehlerwirkung tritt also auf, wenn das Intervall, in welchem interpoliert wird, eine Zeichenfolge, die übrigen Intervalle dagegen lauter Zeichenwechsel aufweisen. Hiernach führt das Differenzenschema der Größen $\operatorname{sg}(U)$ auf eine Tabelle, von der unter Anfügung der Zeilennummern nachstehend ein Ausschnitt hergesetzt ist; die Überschriften geben die Ordnungsnummern an.

	(0)	(1)	(2)	(3)
$a-2$	$+1$		-4	
		-2		$+8$
$a-1$	-1		$+4$	
		$+2$		-6
a	$+1$		-2	
		0		0
$a+1$	$+1$		-2	
		-2		$+6$
$a+2$	-1		$+4$	
		$+2$		-8
$a+3$	$+1$		-4	

In dem vorstehenden Schema ist die Spalte (0) gegen die Zwischenzeile $a + \frac{1}{2}$ symmetrisch angeordnet. Demnach läßt ein Umklappen um diese Zeile die Spalte (0) und folgeweise auch die andern Spalten gerader Ordnung ungeändert, sodaß im besondern für $q = 0, 1, 2, \cdots$

$$(a + 1, 2q) = (a, 2q) \tag{49}$$

wird. Daraus folgt sofort, daß die ungeraden Differenzen auf der Zeile $a + \frac{1}{2}$ sämtlich verschwinden, daß also

$$(a + \tfrac{1}{2}, 2q + 1) = 0 \tag{50}$$

ist.

§ 44. Um die Differenzen zu finden, die in den Formeln N, G, G', G'' gebraucht werden, benutzen wir die in § 11 (3) gefundene Darstellung, wonach die q^{te} Differenz, durch die Funktionswerte ausgedrückt, die Gestalt

$$(a+p-\tfrac{q}{2},q) = (q)_0(a+p,0)-(q)_1(a+p-1,0)+(q)_2(a+p-2,0)-\cdots$$

besitzt. Hiermit erhält man aus dem obigen Differenzenschema zunächst

$$\begin{aligned}
(a+\tfrac{1}{2},1) &= (1)_0-(1)_1,\\
(a+1,2) &= -(2)_0-(2)_1+(2)_2,\\
(a+\tfrac{3}{2},3) &= (3)_0+(3)_1+(3)_2-(3)_3,\\
(a+2,4) &= -(4)_0-(4)_1-(4)_2-(4)_3+(4)_4,
\end{aligned}$$

u. s. w.,

wo das Gesetz des Fortschreitens sich daraus ergibt, daß zwischen den beiden letzten Gliedern rechter Hand jedesmal ein Zeichenwechsel, sonst aber nur Zeichenfolgen auftreten. Beachtet man nun die Relation

$$2^q = (1+1)^q = (q)_0+(q)_1+\cdots+(q)_q,$$

so erhält man den allgemeinen Ausdruck für die vorstehenden Differenzen in der Gestalt

$$(a+\tfrac{1}{2}q,q) = (-1)^q(2-2^q). \tag{51}$$

In derselben Weise erhält man ferner

$$\begin{aligned}
(a,2) &= (2)_0-(2)_1-(2)_2,\\
(a,4) &= -(4)_0-(4)_1+(4)_2+(4)_3+(4)_4,\\
(a,6) &= (6)_0+(6)_1+(6)_2-(6)_3-(6)_4-(6)_5-(6)_6,
\end{aligned}$$

u. s. w.,

wo das Gesetz des Fortschreitens wiederum klar ist, wenn man beachtet, daß die beiden Reihen von Zeichenfolgen rechter Hand durch den einen Zeichenwechsel vor dem mittelsten Gliede getrennt sind. Da nun die symmetrisch zur Mitte gelegenen Glieder sich paarweise aufheben, so ergibt sich der allgemeine Ausdruck

$$(a,2q) = (-1)^q(2q)_q. \tag{52}$$

Hieraus folgt wegen der Symmetrie des Schemas gegen die Zeile $a+\frac{1}{2}$ sofort die Relation

$$(a+1,2q) = (-1)^q(2q)_q. \tag{53}$$

Setzt man schließlich noch die Differenzenbeziehung

$$(a+\tfrac{1}{2},2q+1)-(a-\tfrac{1}{2},2q+1) = (a,2q+2)$$

an, so liefert die Verbindung mit (50) und (53) den Ausdruck

$$(a - \tfrac{1}{2}, 2q + 1) = (-1)^q (2q + 2)_{q+1}.$$

Damit sind die erforderlichen Differenzen hergestellt.

Geht man mit den gefundenen Differenzen in die Gleichungen (27), (30), (31), (32) ein, die zu den Formeln N, G, G', G'' gehören, so wird, wenn man der Größe P jedesmal das betreffende Zeichen N, G, G', G'' anfügt, zunächst die Reihe

$$P(N) = 1 - (2 - 2^1)(t)_1 + (2 - 2^2)(t)_2 - \cdots \tag{54}$$

erhalten. Ferner wird, weil in (30) und (32) jetzt die geraden Glieder übereinstimmen, die ungeraden hingegen wegen (50) herausfallen,

$$P(G) = P(G'') = 1 - (2)_1 (t)_2 + (4)_2 (t+1)_4 - (6)_3 (t+2)_6 - \cdots. \tag{55}$$

Endlich erhält man aus (31)

$$\begin{aligned} P(G') = 1 &+ (2)_1 (t)_1 - (2)_1 (t+1)_2 \\ &- (4)_2 (t+1)_3 + (4)_2 (t+2)_4 \\ &+ (6)_3 (t+2)_5 - (6)_3 (t+3)_6 - \cdots. \end{aligned} \tag{56}$$

Die vorstehenden, nunmehr weiter zu untersuchenden Reihen sind jedesmal mit dem von t abhängenden Binomialkoeffizienten abzubrechen, der den Index n führt.

§ 45. Die Glieder der Reihe für $P(N)$ sind unter der über die Phase t gemachten Voraussetzung sämtlich positiv, folglich wächst $P(N)$ mit wachsendem n. Schreibt man (54) in der Gestalt

$$P(N) = \sum_q (-1)^q (2 - 2^q)(t)_q$$

und bildet den Quotienten aus den Gliedern mit den Nummern $q + 1$ und q, so erhält man den Ausdruck

$$[(2^q - 1)(q - t)] : [(2^{q-1} - 1)(q + 1)],$$

der mit wachsendem q gegen 2 konvergiert. Die Zunahme von $P(N)$ erfolgt also bei größeren Werten von n sehr rasch, da dann jedes neu hinzutretende Glied rund doppelt so groß ist, wie das unmittelbar vorhergehende.

Die Reihe (55) enthält ebenfalls nur positive Glieder. Da jedoch die ungeraden Glieder fehlen, so bleibt $P(G)$ ungeändert, wenn n von $2q$ auf $2q + 1$ übergeht, wogegen bei dem Übergange von $2q + 1$ auf $2q + 2$ ein Zuwachs eintritt. Um zu erkennen, ob $P(G)$ über alle Grenzen wachsen kann, ziehen wir einen Satz aus der *Gauß*schen Abhandlung „Disquisitiones generales circa seriem infinitam etc. (Gesammelte Werke Band III)“ heran. *Gauß* führt daselbst die hypergeometrische Reihe

$$F(a, b, c, x) = 1 + \frac{a}{1}\frac{b}{c}x + \frac{a(a+1)}{1 \cdot 2}\frac{b(b+1)}{c(c+1)}x^2 + \cdots$$

ein, in welcher der Quotient der Glieder der Ordnung $q+1$ und q die Gestalt

$$[(q+a)(q+b)x]:[(q+1)(q+c)] \tag{57}$$

besitzt. Er zeigt ferner, daß für $x=1$ die Glieder der Reihe, wenn $a+b=c$ ist, gegen Null konvergieren, daß jedoch trotzdem die Summe der Reihe unendlich wird.

Bildet man nun in der Reihe (55), die wir in der Gestalt

$$P(G)=\sum_q (-1)^q (2q)_q (t+q-1)_{2q}$$

schreiben können, den Quotienten aus den Gliedern mit den Nummern $q+1$ und q, so entsteht der Ausdruck

$$[(q+t)(q+1-t)]:[(q+1)(q+1)],$$

der mit (57) verglichen zu der Gleichung

$$P(G)=F(t,\, 1-t,\, 1,\, 1)$$

führt. Da hierin $a+b=c$ ist, so geht $P(G)$ mit wachsendem n schließlich über alle Grenzen.

In der Reihe (56) für $P(G')$ sind die ungeraden Glieder positiv, die geraden — von dem Anfangsgliede 1 abgesehen — dagegen negativ. Zieht man ferner jedes ungerade Glied mit dem nachfolgenden geraden vermittels der Beziehung

$$(t+q-1)_{2q-1}-(t+q)_{2q}=-(t+q-1)_{2q}$$

zusammen, so erhält man wieder die Reihe (55), d. h. für ein gerades n stimmt $P(G')$ mit $P(G)$ überein. Geht man bei $P(G')$ von $n=2q$ über auf $n=2q+1$, so erfolgt eine Zunahme, daran schließt sich, wenn n von $2q+1$ auf $2q+2$ übergeht, eine Abnahme, die jedoch kleiner ist als die vorhergehende Zunahme, da ja bei dem Übergange von $n=2q$ auf $n=2q+2$ wegen der hierbei bestehenden Gleichung $P(G')=P(G)$ eine Zunahme stattfindet.

Da die Formeln S und B durch die Mittelung von G und G', sowie von G und G'' entstehen, und da ferner die Größe P für die Formeln G, G', G'' jedesmal auf demselben Fehlersystem beruht, so erhält man $P(S)$ und $P(B)$, wenn man einerseits $P(G)$ und $P(G')$, andererseits $P(G)$ und $P(G'')$ mittelt. Demnach fallen $P(G)$ und $P(B)$ zusammen, ferner fällt für ein gerades n auch $P(S)$ mit $P(G)$ und $P(B)$ zusammen.

§ 46. Eine Vorstellung von dem Betrage der extremen Fehlervergrößerung P für die gewöhnlich vorkommenden Werte von n erhält man aus den nachstehenden Täfelchen, deren Anordnung unmittelbar verständlich ist. Die Phase 0.0 ist fortgelassen, weil für diese die P-Größen sämtlich gleich Eins sind.

Werte von $P(N)$.

$n =$	1	2	3	4	5	6
$t =$ 0.1	1.00	1.09	1.26	1.55	2.03	2.85
0.2	1.00	1.16	1.45	1.92	2.68	3.95
0.3	1.00	1.21	1.57	2.13	3.02	4.46
0.4	1.00	1.24	1.62	2.21	3.10	4.53
0.5	1.00	1.25	1.62	2.17	2.99	4.26
0.6	1.00	1.24	1.58	2.05	2.73	3.77
0.7	1.00	1.21	1.48	1.85	2.37	3.13
0.8	1.00	1.16	1.35	1.60	1.94	2.43
0.9	1.00	1.09	1.18	1.31	1.47	1.70
1.0	1.00	1.00	1.00	1.00	1.00	1.00

Werte von $P(G) = P(B)$.

$n =$	1	2	3	4	5	6
$t =$ 0.1	1.00	1.09	1.09	1.14	1.14	1.17
0.2	1.00	1.16	1.16	1.25	1.25	1.31
0.3	1.00	1.21	1.21	1.33	1.33	1.41
0.4	1.00	1.24	1.24	1.37	1.37	1.47
0.5	1.00	1.25	1.25	1.39	1.39	1.49
0.6	1.00	1.24	1.24	1.37	1.37	1.47
0.7	1.00	1.21	1.21	1.33	1.33	1.41
0.8	1.00	1.16	1.16	1.25	1.25	1.31
0.9	1.00	1.09	1.09	1.14	1.14	1.17
1.0	1.00	1.00	1.00	1.00	1.00	1.00

Werte von $P(S)$.

$n =$	1	2	3	4	5	6
$t =$ 0.1	1.10	1.09	1.14	1.14	1.17	1.17
0.2	1.20	1.16	1.26	1.25	1.31	1.31
0.3	1.30	1.21	1.35	1.33	1.41	1.41
0.4	1.40	1.24	1.41	1.37	1.48	1.47
0.5	1.50	1.25	1.44	1.39	1.51	1.49
0.6	1.60	1.24	1.43	1.37	1.49	1.47
0.7	1.70	1.21	1.39	1.33	1.43	1.41
0.8	1.80	1.16	1.30	1.25	1.33	1.31
0.9	1.90	1.09	1.18	1.14	1.18	1.17
1.0	2.00	1.00	1.00	1.00	1.00	1.00

Nach Ausweis der vorstehenden Zahlen darf für die betrachteten Werte von n die extreme Fehlervergrößerung bei den Formeln G, S, B als unerheblich angesehen werden, namentlich wenn man berücksichtigt, daß die Formel S wesentlich für Phasenwerte zwischen den Grenzen $\pm$ 0.5 bestimmt ist. Ungünstiger liegt die Sache bei der Formel N, indem sich hier von $n = 5$ ab bereits die rasche Divergenz der Reihe für $P(N)$ geltend macht. Es ist das ein neuer Beleg für

die frühere Bemerkung, daß die Formel N weniger vorteilhaft ist, als die Formeln G, S, B.

Bei der Entwicklung des Ausdruckes P war vorausgesetzt worden, daß die für die eigentliche Interpolation nötige Rechnung abrundungsfrei durchgeführt wird. Treten nun hierbei ebenfalls Abrundungen auf, wie das ja in Wirklichkeit stets geschieht, so sind dafür natürlich die entsprechenden Zuschläge zu dem Werte von P zu machen.

§ 47. Zum Schlusse möge noch kurz die Frage berührt werden, wie weit man beim Interpolieren in der Berücksichtigung der Differenzen höherer Ordnung jedesmal zu gehen habe. Die gewöhnliche Antwort lautet: Soweit als die Differenzen einen merklichen Einfluß auf die letzte bei der Rechnung mitzunehmende Decimale ausüben. Diese Regel ist im allgemeinen richtig, es ist jedoch zu beachten, daß das angegebene Kriterium *für sich allein* nicht ausreicht. Bedeutet z. B. $g(x)$ eine ganze Funktion n^{ter} Ordnung und setzt man

$$f(x) = g(x) + \sin 2\pi x,$$

so sind, falls man $f(x)$ mit dem Intervall $h = 1$ tabuliert, die Differenzen oberhalb der Ordnung n in aller Strenge null, trotzdem würde die Interpolation nach den oben entwickelten Formeln im allgemeinen ein stark fehlerhaftes Resultat liefern, weil sie nur die Änderungen von $g(x)$ berücksichtigt. Hieraus ist zu entnehmen, daß die sachgemäße Anwendung der vorliegenden Formeln stets eine bestimmte Kenntnis von dem Verlaufe der Kurve $y = f(x)$ zwischen den durch die Tafel gegebenen Kurvenpunkten voraussetzt. Diese Kenntnis ist selbstverständlich nicht aus der Tafel zu entnehmen, sondern beruht auf dem, was man anderswoher über die Funktion und im besondern über ihre Entwicklung nach dem *Taylor*schen Lehrsatze weiß.

Dritter Abschnitt.

Numerische Differentiation.

§ 48. Die Entwicklungen des § 27 enthielten die Lösung der allgemeinen Interpolationsaufgabe für den Fall, daß die Funktion $f(a+th)$ für ganzzahlige Werte von t tabuliert worden ist. Die Lösung beruhte auf den Formeln

$$f(x) = f(a+th) = \sum_s A_s(a+p_s, s) + K, \tag{1}$$

$$K = R(x) - \sum_s A_s R(p_s, s), \tag{2}$$

$$R(x) = \frac{h^{n+1}}{n!}\int_0^t dv\, v^n f(a+th-vh)_{n+1}. \tag{3}$$

Hierin sind die Größen $(a+p_s, s)$ gewisse aus dem Differenzenschema von $f(x)$ entnommene Glieder, die $R(p_s, s)$ die entsprechenden Glieder aus dem Differenzenschema des Restes $R(x)$, die A_s bestimmte Polynome von t, endlich $f(x)_{n+1}$ die $(n+1)$-ste Ableitung von $f(x)$. Die Korrektion K enthält den Faktor h^{n+1} und läßt sich durch angemessene Verkleinerung von h unmerklich machen, sobald die Ableitung f_{n+1} innerhalb des für die obigen Gleichungen in Betracht kommenden Gebietes der Veränderlichen x endlich bleibt.

Differentiiert oder integriert man die Gleichung (1) nach t, so geht die rechte Seite in einen Ausdruck von ähnlicher Gestalt über. Denn zunächst verwandeln sich die Polynome A_s wieder in Polynome von t. Ferner wird bei der Differentiation

$$dR:dt = \frac{1}{n!} h^{n+1} t^n f(a)_{n+1} + \frac{h^{n+2}}{n!}\int_0^t dv\, v^n f(a+th-vh)_{n+2},$$

d. h. es tritt eine Potenz von t nebst einem ähnlich wie (3) gebauten Integral auf. Andererseits erhält man beim Integrieren von (1) für das Integral von R den nachstehenden, durch Differentiieren leicht zu verifizierenden, Ausdruck

$$\int dt\, R = -\frac{1}{n!} h^n t^n f(a)_n + \frac{h^n}{n!}\int_0^t dv\, v^n f(a+th-vh)_n,$$

d. h. wiederum ein Ergebnis von ähnlicher Art wie vorhin. Hiernach liefert die allgemeine Formel J, die zunächst als Annäherung für die Berechnung von $f(a + th)$ entwickelt worden war, ohne weiteres auch Näherungen für die Ableitung und das Integral von $f(x)$; ebenso sind auch die Voraussetzungen für die Zulässigkeit der neuen Formeln nicht wesentlich von den Bedingungen verschieden, die wir früher bei der Formel J kennen gelernt haben.

Die Rechenvorschriften, die auf dem angegebenen Wege entstehen, werden als „*numerische Differentiation*" und „*numerische Integration*" bezeichnet.

Die numerische Differentiation, der wir uns zunächst zuwenden, besitzt ein verhältnismäßig beschränktes Anwendungsgebiet, weil man es gewöhnlich vorzieht, bei der Berechnung der Ableitung $f(x)_1$ einer analytisch gegebenen Funktion $f(x)$ den analytischen Ausdruck für $f(x)_1$ wirklich herzustellen und danach zu rechnen. Es gibt jedoch bestimmte Fälle, wo die numerische Differentiation sich als sehr nützlich erweist. Es sei z. B. der analytische Ausdruck für $f(x)$ und ebenso für $f(x)_1$ aus einer großen Menge von einzelnen Gliedern zusammengesetzt, sodaß die direkte Berechnung dieser Ausdrücke zeitraubend ist; ferner werde $f(x)$ sehr häufig, $f(x)_1$ hingegen nur selten gebraucht. Dann wird man $f(x)$ ohne Bedenken tabulieren, $f(x)_1$ dagegen nicht. Demgemäß würde man, sobald die Ableitung einmal wirklich gebraucht wird, auf den analytischen Ausdruck dafür zurückgreifen müssen, wenn nicht eben die numerische Differentiation das Mittel böte, die umständliche direkte Berechnung von $f(x)_1$ durch geeignete Benutzung der für $f(x)$ vorhandenen Tafel zu umgehen. Fälle dieser Art liefern unter anderem die sphärischen Koordinaten der Gestirnsörter, die Jahr für Jahr in den Sammlungen astronomischer Ephemeriden vorausberechnet und veröffentlicht werden.

§ 49. Bei der Herleitung der Gebrauchsformeln für die numerische Differentiation beschränken wir uns auf die erste und zweite Ableitung, die wie bisher mit $f(x)_1$ und $f(x)_2$ bezeichnet werden sollen. Ferner sollen nur die Formeln N, S, B herangezogen werden, da die Ausdrücke, die sich aus der Formel G ergeben, keine besondern Vorteile bieten und deshalb bisher bei den Rechnern nicht weiter in Gebrauch gekommen sind.

Die Formel N besitzt die Gestalt

$$f(a \pm th) = A_0(a, 0) \pm A_1(a \pm \tfrac{1}{2}, 1) + A_2(a \pm 1, 2) \pm \cdots,$$

worin A_q mit dem Binomialkoeffizienten $(t)_q$ zusammenfällt. Bezeichnet man die erste und zweite Ableitung eines A durch die Buchstaben B und C, so wird, weil $A_0 = 1$ und $A_1 = t$ ist,

$$hf(a \pm th)_1 = B_1(a \pm \tfrac{1}{2}, 1) \pm B_2(a \pm 1, 2) + B_3(a \pm \tfrac{3}{2}, 3) + \cdots, \quad (4)$$

$$h^2 f(a \pm th)_2 = C_2(a \pm 1, 2) \pm C_3(a \pm \tfrac{3}{2}, 3) + C_4(a \pm 2, 4) + \cdots \quad (5)$$

Bezeichnet man mit b, c die Werte, welche die Polynome B, C für $t = 0$ annehmen, so entsehen aus (4) und (5) mit $t = 0$ noch die Formeln

$$hf(a)_1 = b_1(a \pm \tfrac{1}{2}, 1) \pm b_2(a \pm 1, 2) + \cdots, \quad (6)$$

$$hf(a)_2 = c_2(a \pm 1, 2) \pm c_3(a \pm \tfrac{3}{2}, 3) + \cdots. \quad (7)$$

Um die Koeffizienten zu ermitteln setze man für den Augenblick

$$f(x) = \exp x, \quad S = \exp(\tfrac{1}{2}h) - \exp(-\tfrac{1}{2}h),$$

$$S \exp(\tfrac{1}{2}h) = k, \quad \exp h = 1 + k, \quad h = \log(1 + k).$$

Dann ist nach § 17 $(a + p, q) = \exp(a + ph) S^q$, woraus

$$(a + \tfrac{1}{2}q, q) = \exp(a + \tfrac{1}{2}qh) S^q = \exp a\, k^q$$

folgt. Da ferner jetzt

$$f(a + th)_1 = f(a + th)_2 = \exp(a + th) = \exp a\,(1 + k)^t$$

wird, so entstehen aus (4) bis (7) die Gleichungen

$$h\,(1 + k)^t = B_1 k + B_2 k^2 + \cdots,$$

$$h^2(1 + k)^t = C_2 k^2 + C_3 k^3 + \cdots,$$

$$h = b_1 k + b_2 k^2 + \cdots,$$

$$h^2 = c_2 k^2 + c_3 k^3 + \cdots.$$

Entwickelt man $h = \log(1 + k)$ nach k, so ergeben sich die Koeffizienten b; durch Quadrieren der Reihe findet man die Größen c. Werden ferner diese beiden Reihen mit der Reihe

$$(1 + k)^t = (t)_0 + (t)_1 k + (t)_2 k^2 + \cdots$$

multipliziert, so erhält man die Koeffizienten B und C. Dem entsprechend läßt sich zunächst folgendes Täfelchen aufstellen:

$$\left.\begin{array}{lcccccc} q = & 1 & 2 & 3 & 4 & 5 & 6 \\ b = & 1 & -\frac{1}{2} & \frac{1}{3} & -\frac{1}{4} & \frac{1}{5} & -\frac{1}{6} \\ c = & - & 1 & -1 & \frac{11}{12} & -\frac{5}{6} & \frac{137}{180} \end{array}\right\}. \quad (8)$$

Ferner wird

$$\left.\begin{array}{l} B_1 = b_1(t)_0, \\ B_2 = b_2(t)_0 + b_1(t)_1, \\ B_3 = b_3(t)_0 + b_2(t)_2 + b_1(t)_2, \end{array}\right\} \quad (9)$$

u. s. w.

und ebenso

$$\left.\begin{aligned} C_2 &= c_2(t)_0, \\ C_3 &= c_3(t)_0 + c_2(t)_1, \\ C_4 &= c_4(t)_0 + c_3(t)_1 + c_2(t)_2, \\ &\text{u. s. w.} \end{aligned}\right| \qquad (10)$$

Die Rechnung nach den vorstehenden Gleichungen wird, namentlich wenn man Tafeln für die Koeffizienten $(t)_q$ zur Hand hat, etwas übersichtlicher, als wenn man die Ausdrücke B, C nach Potenzen von t entwickelt hat. Im übrigen gilt die schon gelegentlich der Interpolation gemachte Bemerkung, daß die Formel N nur als Aushilfe für diejenigen Fälle zweckmäßig ist, in denen man es mit dem Anfang oder Schluß des Differenzenschemas zu tun hat; es wird sich zeigen, daß die Formeln S und B zu sehr viel bequemeren Ausdrücken für die numerische Differentiation führen.

§ 50. Die Formel S liefert nach § 32 (33) eine Reihe von der Gestalt

$$f(a + th) = A_0(a, 0) + A_1(a, 1) + A_2(a, 2) + \cdots,$$

worin

$$\begin{aligned} A_0 &= 1, & A_1 &= t, & 2A_2 &= t^2, \\ 6A_3 &= t^3 - t, & & & 24A_4 &= t^4 - t^2, \\ 120A_5 &= t^5 - 5t^3 + 4t, & & & 720A_6 &= t^6 - 5t^4 + 4t^2, \\ &\text{u. s. w.} \end{aligned}$$

ist. Die Differentiation liefert die Reihen

$$hf(a + th)_1 = B_1(a, 1) + B_2(a, 2) + \cdots, \qquad (11)$$

$$h^2 f(a + th)_2 = C_2(a, 2) + C_3(a, 3) + \cdots, \qquad (12)$$

nebst den Ausdrücken

$$\begin{aligned} B_1 &= 1, \\ B_2 &= t, \\ 6B_3 &= 3t^2 - 1, \\ 12B_4 &= 2t^3 - t, \\ 120B_5 &= 5t^4 - 15t^2 + 4, \\ 360B_6 &= 3t^5 - 10t^3 + 4t, \\ &\text{u. s. w.} \end{aligned} \qquad (13)$$

$$\begin{aligned} C_2 &= 1, \\ C_3 &= t, \\ 12C_4 &= 6t^2 - 1, \\ 12C_5 &= 2t^3 - 3t, \\ 360C_6 &= 15t^4 - 30t^2 + 4, \\ &\text{u. s. w.} \end{aligned} \qquad (14)$$

Daraus folgt für $t = 0$

$$hf(a)_1 = (a, 1) - \tfrac{1}{6}(a, 3) + \tfrac{1}{30}(a, 5) - \cdots \tag{15}$$

$$h^2 f(a)_2 = (a, 2) - \tfrac{1}{12}(a, 4) + \tfrac{1}{90}(a, 6) - \cdots. \tag{16}$$

Hierin treten nur gerade oder nur ungerade Glieder auf, weil die Koeffizienten A abwechselnd gerade und ungerade Funktionen von t sind.

In der für die Formel B geltenden Reihe

$$f(a + th) = A_0(a + \tfrac{1}{2}, 0) + A_1(a + \tfrac{1}{2}, 1) + A_2(a + \tfrac{1}{2}, 2) + \cdots$$

sind die Koeffizienten abwechselnd gerade und ungerade Funktionen der Größe $u = t - \frac{1}{2}$. Entwickelt man demgemäß die in § 32 (34) aufgestellten Koeffizienten nicht nach t, sondern nach u, so wird

$$A_0 = 1, \qquad A_1 = u, \qquad 8A_2 = 4u^2 - 1,$$

$$24A_3 = 4u^3 - u, \qquad 384A_4 = 16u^4 - 40u^2 + 9,$$

$$1920A_5 = 16u^5 - 40u^3 + 9u,$$

$$46080A_6 = 64u^6 - 560u^4 + 1036u^2 - 225,$$

u. s. w.

Hieraus entspringen die Reihen

$$hf(a + th)_1 = B_1(a + \tfrac{1}{2}, 1) + B_2(a + \tfrac{1}{2}, 2) + \cdots, \tag{17}$$

$$h^2 f(a + th)_2 = C_2(a + \tfrac{1}{2}, 2) + C_3(a + \tfrac{1}{2}, 3) + \cdots, \tag{18}$$

nebst den Ausdrücken

$$\begin{aligned} B_1 &= 1, \\ B_2 &= u, \\ 24B_3 &= 12u^2 - 1, \\ 24B_4 &= 4u^3 - 5u, \\ 1920B_5 &= 80u^4 - 120u^2 + 9, \\ 5760B_6 &= 48u^5 - 280u^3 + 259u, \end{aligned} \tag{19}$$

u. s. w.

und weiter

$$\begin{aligned} C_2 &= 1, \\ C_3 &= u, \\ 24C_4 &= 12u^2 - 5, \\ 24C_5 &= 4u^3 - 3u, \\ 5760C_6 &= 240u^4 - 840u^2 + 259, \end{aligned} \tag{20}$$

u. s. w.

Setzt man $t = \frac{1}{2}$, also $u = 0$, so entstehen die Formeln für die „*Differentiation in die Mitte*", nämlich

$$hf(a + \tfrac{1}{2}h)_1 = (a + \tfrac{1}{2}, 1) - \tfrac{1}{24}(a + \tfrac{1}{2}, 3) + \tfrac{3}{640}(a + \tfrac{1}{2}, 5) - \cdots \qquad (21)$$

$$h^2 f(a + \tfrac{1}{2}h)_2 = (a + \tfrac{1}{2}, 2) - \tfrac{5}{24}(a + \tfrac{1}{2}, 4) + \tfrac{259}{5760}(a + \tfrac{1}{2}, 6) - \cdots \qquad (22)$$

Hierin treten wieder nur gerade und nur ungerade Glieder auf.

§ 51. Für die in (13), (14), (19), (20) angesetzten Koeffizienten B, C findet man Tafeln in der oben § 36 genannten Sammlung von *Bauschinger.* Wenn keine Tafeln zur Verfügung stehen, so wird die Rechnung übersichtlicher, indem man nicht die für ein beliebiges Argument $a + th$ geltenden Reihen anwendet, sondern unter Benutzung der Formeln (15), (16), (21), (22) für die gesuchten Ableitungen ein kleines Täfelchen anlegt und darin interpoliert.

Würde man die oben entwickelten Formeln auf das in § 15 als Tabelle IV gegebene Täfelchen anwenden, so erhielte man offenbar die erste Ableitung von Log x mit fünf, die zweite mit vier geltenden Ziffern. Legte man dagegen eine Tafel mit zehnmal kleinerem Intervall zu Grunde, so enthielte die erste Ableitung nur noch vier, die zweite nur noch zwei geltende Ziffern. Hiernach darf in dem Differenzenschema, das zur numerischen Differentiation dienen soll, das Intervall h nicht zu klein gewählt werden. Andererseits darf man aber auch mit dem Intervall nicht über eine gewisse obere Grenze hinausgehen, da mit der Vergrößerung des Intervalls die Anzahl der mitzunehmenden Differenzen wächst, was wiederum Nachteile anderer Art im Gefolge hat. Demnach wird für die Intervallgröße jedesmal ein gewisses Optimum vorhanden sein, für dessen Auffindung sich allerdings keine allgemeinen Regeln angeben lassen.

Will man dem Bildungsgesetz der Koeffizienten in den oben entwickelten Reihen nachgehen, so kann man hierzu die Festsetzung $f(x) = \exp x$ benutzen. Da jedoch diese Untersuchung ebenso wie die gleiche bei der numerischen Integration auftretende Aufgabe für die gewöhnlichen Bedürfnisse des wissenschaftlichen Rechnens nur von geringer Bedeutung ist, so begnüge ich mich damit, auf die Behandlung hinzuweisen, die der Gegenstand in den nachstehend genannten Werken gefunden hat:

P. A. Hansen. Relationen einesteils zwischen Summen und Differenzen und anderenteils zwischen Integralen und Differentialen. (Abhandlungen der mathematisch-physischen Klasse der K. Sächs. Gesellschaft der Wissenschaften, VII. Band 1865.)

Th. von Oppolzer. Lehrbuch zur Bahnbestimmung der Kometen und Planeten. Zweiter Band, Leipzig 1880.

Vierter Abschnitt.

Numerische Integration: Summenmethode.

§ 52. Die Ausführung einer Integration wird häufig auch als Quadratur bezeichnet, weil sich der Wert eines bestimmten Integrals geometrisch als der Flächeninhalt einer gewissen Figur deuten läßt und weil ferner die geforderte Flächenbestimmung darauf hinauskommt, ein Quadrat zu finden, das mit der vorgelegten Figur flächengleich ist. Des weitern findet man in der Literatur für die rein numerisch ausgeführte Integration seit langer Zeit den Namen „mechanische Quadratur" gebraucht, der einen Gegensatz zu der mit Zirkel und Lineal konstruierenden „geometrischen" Quadratur ausdrücken soll. Ein solcher Gegensatz war früher in der Tat vorhanden, denn die ältere Geometrie behandelte ihre Gebilde als starr, während die Analysis schon auf den ersten Stufen ihrer Entwicklung davon ausging, die betrachteten Gebilde als fließend, als durch Bewegung entstanden aufzufassen. Heute ist jene alte Gegenüberstellung von „geometrisch" und „mechanisch" bedeutungslos geworden, denn die neuere Geometrie macht ebenfalls in ausgedehnter Weise Gebrauch von der Vorstellung fließender Gebilde. Diesem Umstande sollte man meines Erachtens jetzt auch bei der Namengebung Rechnung tragen: der alte Name Quadratur mag als generelle Bezeichnung neben dem Worte Integration bestehen bleiben, das im letzten Grunde die Summation infinitesimaler Elemente ausdrückt; im übrigen unterscheide man bis auf weiteres zwischen einer *analytischen*, einer *numerischen* und einer *mechanischen* Integration, wobei der letzte Name den Methoden zuzuweisen ist, die mit integrierenden Mechanismen arbeiten. Es sind ja schon jetzt verschiedene äußerst sinnreich erdachte Integrationsmaschinen vorhanden, die sich innerhalb bestimmter Verwendungsgebiete als höchst nützlich erwiesen haben: die bekannteste darunter ist das *Amsler*sche Planimeter mit seinen Abkömmlingen.

Für die numerische Integration, der wir uns jetzt zuwenden, war in § 48 bereits auf einen Weg hingewiesen worden, der von den früher gefundenen Interpolationsformeln ausgeht. Dieser Weg ist jedoch, wie sich zeigen wird, nicht der einzige gangbare; je nachdem man für die analytische Darstellung einer vorläufig unbestimmt gelassenen

Funktion $f(x)$ andere und andere Formen zu Grunde legt, findet man auch andere und andere Vorschriften für die Ausführung der numerischen Integration.

Die Methoden, von denen nachstehend die Rede sein wird, zerfallen in zwei Gruppen, je nachdem bei der Auswertung eines vorgelegten Integrals die Werte der zu integrierenden Funktion nur *innerhalb* des Integrationsgebietes oder aber auch *außerhalb* desselben herangezogen werden. Die erste Gruppe umfaßt die Methoden der „*Viereckverbesserung*“ und der „*Mittelwerte*“, die zweite Gruppe beruht dagegen auf der Benutzung der Summenreihen und soll kurz als die „*Summenmethode*“ bezeichnet werden. Da die Summenmethode unmittelbar mit der Interpolation zusammenhängt, so möge ihre Herleitung an die Spitze gestellt werden.

§ 53. Um das Integral einer tabulierten Funktion $f(x)$ aus den Tafelwerten zu finden, werden von den Rechnern — man darf wohl sagen ausschließlich — die Formeln benutzt, die sich ähnlich wie die Interpolationsformeln S und B aus den Gliedern derselben Haupt- oder Zwischenzeile zusammensetzen, weil andere Ausdrücke, die man ebenfalls ohne Schwierigkeit bilden könnte, weniger bequem sind. Bei der Herleitung gehen wir von der Formel S aus; ferner ziehen wir sogleich den Fall der wiederholten Integration mit heran und schreiben

$$\int f(x)\,dx = J(x)_1, \qquad \int J(x)_1\,dx = J(x)_2, \quad \text{u. s. w.},$$

wo die willkürlichen Integrationskonstanten irgendwie bestimmt sein mögen. Wird der reziproke Wert des Tafelintervalls h mit k bezeichnet, so ist $k = 1 : h$ und weiter

$$kJ(a + th)_1 = \int f(a + th)\,dt. \tag{1}$$

Hierin ist jetzt für $f(a + th)$ die durch die Formel S gegebene Reihe einzuführen, wobei die Korrektion K außer Betracht gelassen werden soll.

Schreibt man die Formel S in der Gestalt

$$f(a + th) = A_0(a, 0) + A_1(a, 1) + A_2(a, 2) + \cdots$$

und integriert nach t von $-\frac{1}{2}$ bis $+\frac{1}{2}$, so erhält man linker Hand wegen (1) den Ausdruck

$$kJ(a + \tfrac{1}{2}\,h)_1 - kJ(a - \tfrac{1}{2}\,h)_1.$$

Rechter Hand fallen die ungeraden Glieder heraus, weil die zugehörigen Koeffizienten A ungerade Funktionen von t sind. Ferner liefern die geraden Koeffizienten $A_{2\nu}$ durch die Integration bestimmte Zahlen-

werte b_q, sodaß, wenn man auf beiden Seiten mit h multipliziert, die Gleichung

$$J(a + \tfrac{1}{2} h)_1 - J(a - \tfrac{1}{2} h)_1 = \sum_q b_q h(a, 2q) \tag{2}$$

entsteht, in welcher der Summationsindex q, ebenso wie weiterhin, von Null an läuft. Um hieraus die Werte von J_1 zu finden, denke man sich zunächst das Differenzenschema von $f(x)$ durch Hinzufügung der Summenspalten vervollständigt, wobei die auftretenden Summenkonstanten vorläufig willkürlich bleiben sollen. Ferner denke man sich noch für den Augenblick auf jeder Zeile $a + \frac{1}{2}$ einerseits die Größe $J(a + \frac{1}{2} h)_1$, andererseits den Wert des Ausdruckes

$$F(a + \tfrac{1}{2}) = \sum_q b_q h(a + \tfrac{1}{2}, 2q - 1) \tag{3}$$

eingetragen. Dadurch entstehen zwei neue Spalten, deren erste Differenzen offenbar auf die linke und rechte Seite von (2) führen, also Zeile für Zeile übereinstimmen. Hiernach unterscheiden sich die Größen J_1 und F nur um eine von der Zeilennummer unabhängige feste Größe C, deren Wert von den beiden willkürlichen Konstanten abhängt, die dem Integral J_1 und der in (3) auftretenden Summengröße $(a + \frac{1}{2}, -1)$ anhaften. Denkt man sich nun in der Gleichung $J_1 = F + C$ die Größe C zu der in F enthaltenen Summenkonstante geschlagen, so entsteht die Gleichung

$$J(a + \tfrac{1}{2} h)_1 = \sum_q b_q h(a + \tfrac{1}{2}, 2q - 1). \tag{4}$$

Die hierin auftretenden Koeffizienten b_q sollen weiterhin zugleich mit den Koeffizienten für die übrigen noch aufzustellenden Formeln bestimmt werden.

Der Inhalt der Gleichung (4) läßt sich folgendermaßen in Worte kleiden: Wird die Funktion $f(x)$ mit dem Intervall h von der Stelle $x = a$ aus tabuliert und das Differenzenschema dazu gebildet, so erhält man das Integral von $f(x)$ für die Zeilen $a + \frac{1}{2} h$ durch die in (3) angegebene Operation F, die darauf hinauskommt, daß man in dem f-Schema auf der Zeile $a + \frac{1}{2} h$ die Glieder mit den Ordnungsnummern $-1, 1, 3, \cdots$ vermittelst der Multiplikatoren $b_q h$ verbindet.

Die vorstehende Regel führt sofort zu den Werten des zweiten Integrals J_2, wenn man an die Stelle der Funktion $f(x)$ das Integral $J(x + \frac{1}{2} h)_1$ treten läßt. Die geforderte Tabulierung von der Stelle $x = a$ aus läßt sich unmittelbar mit den aus (4) zu berechnenden Werten bewerkstelligen, ferner führen die Summen- und Differenzenoperationen, an der rechten Seite von (4) vorgenommen, für die Ordnungsnummern $-1, 1, 3, \cdots$ und für die Zeile $x = a + \frac{1}{2} h$ zu den Ausdrücken

$$\sum_q b_q h(a + 1, 2q - 2), \quad \sum_q b_q h(a + 1, 2q), \quad \sum_q b_q h(a + 1, 2q + 2), \cdots,$$

welche vermittelst der Multiplikatoren $b_q h$ verbunden die Formel

$$J(a + h)_2 = \sum_q c_q h^2 (a + 1, 2q - 2) \tag{5}$$

liefern, in der die Größen c_q wiederum bestimmte Zahlenfaktoren bedeuten.

Die nochmalige Ausführung des Verfahrens liefert, wie leicht zu sehen, für J_3 eine Darstellung von der Gestalt

$$J(a + \tfrac{3}{2} h)_3 = \sum_q d_q h^3 (a + \tfrac{3}{2}, 2q - 3), \tag{6}$$

und man erkennt, daß allgemein die Werte der durch n-malige Integration von $f(x)$ erzeugten Funktion ausgedrückt werden durch die Summengrößen n^{ter} Ordnung von $f(x)$ und durch deren Differenzen.

§ 54. Die Formeln (4), (5), (6) enthalten immer nur Glieder der zu $f(x)$ gehörigen Haupttabelle. Um die entsprechenden aus den Größen der Zwischentabelle zusammengesetzten Formeln zu erhalten genügt es, die gefundenen Darstellungen in die Mitte zu interpolieren. Bildet man zu der rechten Seite von (4) zunächst die Differenzen der geraden Ordnung $2r$ und daraus dann die Zwischengrößen, so entstehen die Ausdrücke

$$\sum_q b_q h(a, 2q + 2r - 1).$$

Setzt man hierin für r der Reihe nach $0, 1, 2, \cdots$ und verbindet die entstehenden Größen vermittelst der Multiplikatoren

$$+1, \quad -1:8, \quad +3:128, \quad -5:1024, \text{ u. s. w.},$$

die bei der Interpolation in die Mitte auftreten, so wird der Ausdruck für $J(a)_1$ erhalten. Damit ergibt sich eine Darstellung von der Gestalt

$$J(a)_1 = \sum_q e_q h(a, 2q - 1). \tag{7}$$

In ähnlicher Weise entstehen aus (5) und (6) die Reihen

$$J(a + \tfrac{1}{2} h)_2 = \sum_q f_q h^2 (a + \tfrac{1}{2}, 2q - 2), \tag{8}$$

$$J(a + h)_3 = \sum_q g_q h^3 (a + 1, 2q - 3). \tag{9}$$

Nach diesen Entwicklungen sind nunmehr noch die Werte der Koeffizienten $b, c, \cdots$ zu ermitteln.

§ 55. Die auf der Haupttabelle beruhenden Formeln (4), (5), (6) und deren Fortsetzung lassen sich einheitlich in der Gestalt

$$J(a + \tfrac{1}{2} rh)_r = \sum_q r_q h^r (a + \tfrac{1}{2} r, 2q - r) \tag{10}$$

schreiben, wo r der Reihe nach die Werte 1, 2, 3, ·· anzunehmen hat. Ebenso kann man, wenn dem Index s die Werte 0, 1, 2, ·· erteilt werden, die Formeln (7), (8), (9), die zur Zwischentabelle gehören, durch die Gleichung

$$J(a + \tfrac{1}{2}\,sh)_{s+1} = \sum_q s_q h^{s+1}(a + \tfrac{1}{2}\,s,\ 2q - s - 1) \tag{11}$$

zusammenfassen. Wird nun statt $f(x)$ die Exponentialfunktion $\exp x$ eingeführt, so hat man zu setzen

$$S = \exp(\tfrac{1}{2}h) - \exp(-\tfrac{1}{2}h), \qquad 2C = \exp(\tfrac{1}{2}h) + \exp(-\tfrac{1}{2}h),$$
$$(a + p,\ q) = \exp(a + ph)S^q \quad \text{für die Haupttabelle,}$$
$$(a + p,\ q) = \exp(a + ph)CS^q \quad \text{für die Zwischentabelle.}$$

Ferner wird das r^{te} Integral $J(x)_r$ gleich der Funktion $\exp x$ plus einer ganzen Funktion $(r-1)^{\text{sten}}$ Grades, welche die willkürlichen Integrationskonstanten enthält und kurz mit $P(x)_{r-1}$ bezeichnet werden soll. Damit erhält man aus (10) und (11) die Gleichungen

$$\exp(a + \tfrac{1}{2}\,rh) + P(a + \tfrac{1}{2}\,rh)_{r-1} = \sum_q r_q h^r \exp(a + \tfrac{1}{2}\,rh)\,S^{2q-r}$$
$$\exp(a + \tfrac{1}{2}\,sh) + P(a + \tfrac{1}{2}\,sh)_s \quad = \sum_q s_q h^{s+1} \exp(a + \tfrac{1}{2}\,sh)\,CS^{2q-s-1}.$$

Die vorstehenden Gleichungen denke man sich tabuliert, indem man für a der Reihe nach die Argumente der Hauptzeilen von $f(x)$ einsetzt. Ferner suche man zu diesen Tafeln die Differenzen n^{ter} Ordnung auf, wobei n mindestens gleich r oder $s + 1$ sein soll. Dann fallen die von den Integrationskonstanten abhängenden P-Größen heraus, und man erhält

$$\exp(a + \tfrac{1}{2}\,rh + \tfrac{1}{2}\,nh)S^n = \sum_q r_q h^r \exp(a + \tfrac{1}{2}\,rh + \tfrac{1}{2}\,nh)\,S^{2q-r+n}$$
$$\exp(a + \tfrac{1}{2}\,sh + \tfrac{1}{2}\,nh)S^n = \sum_q s_q h^{s+1} \exp(a + \tfrac{1}{2}\,sh + \tfrac{1}{2}\,nh)\,CS^{2q-s+n-1}.$$

Hieraus fließen, wenn man die gemeinsamen Faktoren heraushebt und wieder $k = 1 : h$ einführt, die Reihen

$$(kS)^r = \sum_q r_q S^{2q}, \qquad (kS)^{s+1} = \sum_q s_q CS^{2q}, \tag{12}$$

worin $r = 1, 2, \cdots$ und $s = 0, 1, \cdots$ zu setzen ist.

Die Ausdrücke für S und C geben, wenn man nach h differentiiert,

$$dS = Cdh, \qquad 4dC = Sdh, \tag{13}$$

ferner ist

$$C^2 = 1 + \tfrac{1}{4}S^2. \tag{14}$$

Hieraus folgt zunächst

$$\frac{1}{C} = (1 + \tfrac{1}{4}S^2)^{-\frac{1}{2}} \tag{15}$$
$$= 1 - \tfrac{1}{8}S^2 + \tfrac{3}{128}S^4 - \tfrac{5}{1024}S^6 + \cdots.$$

Multipliziert man beide Seiten mit dS und integriert, so entsteht, weil $dh = dS : C$ ist, für h die Reihe

$$h = S - \tfrac{1}{24} S^3 + \tfrac{3}{640} S^5 - \tfrac{5}{7168} S^7 + \cdots . \qquad (16)$$

Durch direkte Division ergibt sich hieraus der Quotient $h : S = kS$ in der Gestalt

$$kS = 1 + \tfrac{1}{24} S^2 - \tfrac{17}{5760} S^4 + \tfrac{367}{967680} S^6 - \cdots . \qquad (17)$$

Diese Reihe nebst ihren Potenzen liefert die Koeffizienten r_q. Multipliziert man ferner die für die Potenzen von kS gefundenen Reihen noch mit der in (15) angegebenen Reihe für $1 : C$, so gelangt man mit Rücksicht auf (12) zu den Koeffizienten s_q.

Nachstehend sind für die Integralnummern $r = 1, 2, 3$ und für die Differenzennummern n die vier niedrigsten Koeffizienten jeder Formel zusammengestellt.

Koeffizienten in $J(a)_r$

$r =$	1	2	3
$n = -3$	—	—	+ 1
-2	—	+ 1	—
-1	+ 1	—	0
0	—	+ 1 : 12	—
1	− 1 : 12	—	+ 1 : 240
2	—	− 1 : 240	—
3	+ 11 : 720	—	− 31 : 30240
4	—	+ 31 : 60480	—
5	− 191 : 60480	—	

Koeffizienten in $J(a + \tfrac{1}{2} h)_r$

$r =$	1	2	3
$n = -3$	—	—	+ 1
-2	—	+ 1	—
-1	+ 1	—	+ 1 : 8
0	—	− 1 : 24	—
1	+ 1 : 24	—	− 7 : 1920
2	—	+ 17 : 1920	—
3	− 17 : 5760	—	+ 457 : 967680
4	—	− 367 : 193536	—
5	+ 367 : 967680	—	

Nachdem die Formeln gebrauchsfertig aufgestellt sind, ist noch anzugeben, wie die Summenkonstanten gefunden werden.

§ 56. Setzt man in jeder der zu benutzenden Summenspalten je ein Glied beliebig — sagen wir der Einfachheit halber mit einer Null — an, so liefert die Durchrechnung der oben entwickelten Formeln die

völlig bestimmten Werte gewisser Integralfunktionen, die für den Augenblick mit $K(x)_1$, $K(x)_2$, $\cdot\cdot$ bezeichnet werden mögen. Die zu einem anderen Konstantensystem gehörigen Integrale $J(x)_1$, $J(x)_2$, $\cdot\cdot$ sind dann mit den $K(x)$ durch Gleichungen der Form

$$\begin{aligned} J(x)_1 &= K(x)_1 + A, \\ J(x)_2 &= K(x)_2 + Ax + B, \\ J(x)_3 &= K(x)_3 + \tfrac{1}{2} Ax^2 + Bx + C, \\ &\text{u. s. w.} \end{aligned} \tag{18}$$

verbunden, in denen die $A, B, C, \cdot\cdot$ Integrationskonstanten bedeuten. Wird nun, wie das meistens geschieht, zur Bestimmung der Integrationskonstanten vorgeschrieben, daß die $J(x)$ für ein vorgeschriebenes Argument $x = X$ vorgeschriebene Werte L annehmen sollen, so entstehen die Gleichungen

$$\begin{aligned} L_1 &= K(X)_1 + A, \\ L_2 &= K(X)_2 + AX + B, \\ &\text{u. s. w.} \end{aligned} \tag{19}$$

in denen die $K(X)$ aus der vorliegenden Tafel der Größen $K(x)$ entweder unmittelbar auszuschreiben oder durch gewöhnliche Interpolation zu entnehmen sind. Hieraus ergeben sich dann die Konstanten $A, B, \cdot\cdot$, sodaß nunmehr auch die gesuchten $J(x)$ ziffermäßig völlig bestimmt sind.

Für die Berechnung der $J(x)$ würde es offenbar eine Erleichterung bedeuten, wenn man die Konstanten $A, B, \cdot\cdot$ von vornherein auf Null bringen könnte. Dies läßt sich durch eine kleine Vorbereitung erreichen. Man setzt in der ersten Summenspalte auf der zur Haupttabelle gehörigen Zeile, die der Stelle $x = X$ anliegt, zunächst eine Null an und berechnet die $K(x)_1$ für einige Zeilen, welche die Stelle $x = X$ einschließen. Hierauf interpoliert man den Wert von $K(X)_1$ und ermittelt nach (19) die Größe A. Danach wiederholt man die Rechnung, indem man statt der vorher benutzten Null den für A gefundenen Wert als Ausgangsglied der Rechnung einträgt und nunmehr die erste Summenspalte samt den daraus folgenden Integralwerten fertig rechnet. Da die solchergestalt gefundenen Integralwerte um den Betrag A größer sind, als die vorher ermittelten $K(x)_1$, so erhält man unmittelbar die gesuchten Größen $J(x)_1$. In derselben Weise verfährt man bei der Ermittelung der höheren Integrale.

Um die beschriebene Vorbereitung entbehrlich zu machen, wird von den Rechnern gern die Bedingung eingeführt, daß die $J(x)$ auf einer bestimmten Haupt- oder Zwischenzeile verschwinden sollen.

Bedeutet e je nach Umständen den Wert o oder $\frac{1}{2}$, so läßt sich das Argument dieser „Nullzeile“ in der Form $b + eh$ schreiben, wo b ein rundes Tafelargument ist. Hiermit nimmt die Nullbedingung zunächst die Gestalt

$$o = J(b + eh)_1 = h(b + e, -1) + c_1 h(b + e, 1) + \cdots$$

an, worin die Koeffizienten $c_1, \cdots$ aus der oben gegebenen Zusammenstellung zu entnehmen sind. Die vorstehende Reihe enthält außer dem Anfangsglied, das für gewöhnlich den Hauptbeitrag zu dem gesuchten Integral liefert, noch die übrigen mit kleinen Faktoren behafteten Terme, deren Gesamtbetrag wir kurz als die „Verbesserung“ des Anfangsgliedes bezeichnen wollen. Das Verschwinden des Integrals besagt nun, daß das Anfangsglied gleich der mit umgekehrten Vorzeichen zu nehmenden Verbesserung sein soll, deren Betrag unmittelbar aus den bereits vorhandenen Zahlen des Differenzenschemas zu entnehmen ist. Damit ist offenbar der Wert von $(b + e, -1)$ festgelegt. Gehört nun diese Größe der Haupttabelle an, so hat man nur nötig, die gefundene Zahl in die erste Summenspalte einzutragen und die Spalte fertig zu rechnen. Ist dagegen $(b + e, -1)$ eine Zwischengröße, so findet man die beiden anliegenden Hauptgrößen durch Addition und Subtraktion der beiden Gleichungen

$$\begin{aligned} \tfrac{1}{2}(b + e + \tfrac{1}{2}, -1) + \tfrac{1}{2}(b + e - \tfrac{1}{2}, -1) &= (b + e, -1), \\ \tfrac{1}{2}(b + e + \tfrac{1}{2}, -1) - \tfrac{1}{2}(b + e - \tfrac{1}{2}, -1) &= \tfrac{1}{2}(b + e, 0), \end{aligned} \qquad (20)$$

deren rechte Seiten bekannt sind. Mit den so ermittelten Zahlen hat man dann wieder die erste Summenspalte fertig zu rechnen.

Die vorstehende Regel, nämlich „*Anfangsglied gleich der umgekehrten Verbesserung*“ gilt offenbar auch für die Integrale der zweiten, dritten u. s. w. Ordnung.

In den Reihen für $J(x)_r$ enthalten die einzelnen Glieder den Faktor h^r. Die besondere Berücksichtigung dieses Faktors wird umgangen, wenn man bei der Anlegung des Differenzenschemas nicht die Werte von $f(x)$, sondern sogleich von $h^r f(x)$ tabuliert, denn dann enthalten sämtliche Zahlen des Schemas bereits den Faktor h^r. Hierbei ist nur zu beachten, daß die Rechnung die Integrale $J_1, J_2, \cdots$ mit den Faktoren $h^{r-1}, h^{r-2}, \cdots$ behaftet liefert.

Tabuliert man nicht $h^r f(x)$ sondern $\frac{1}{2} h^r f(x)$, so ergeben sich die richtigen Zwischengrößen, wenn man die benachbarten Hauptgrößen nicht mittelt, sondern summiert. Hierin liegt eine gewisse Erleichterung, wenn man hauptsächlich mit Zwischengrößen rechnen will.

§ 57. Zum besseren Verständnis der späteren Bemerkungen möge jetzt ein Beispiel behandelt werden. Hierzu soll dasselbe Inte-

gral dienen, das *Gauß* bei der Darlegung der von ihm gefundenen Mittelwertmethode benutzt hat, nämlich das zwischen den Grenzen 100000 und 200000 genommene Integral

$$\int \frac{dx}{\log x}. \tag{21}$$

Demnach ist, wenn M den Modulus der gemeinen Logarithmen bedeutet, der Ausdruck

$$h f(x) = h M : \operatorname{Log} x \tag{22}$$

zu tabulieren und zu dem Ende vorerst über das Intervall h Verfügung zu treffen. Nun beruht das ganze Verfahren darauf, daß man die Interpolationsformel S oder die Gleichung

$$f(a + th) = A_0(a, 0) + \cdot \cdot + A_n(a, n) + K$$

unter Vernachlässigung der Korrektion K nach t zwischen den Grenzen $t = \pm 0.5$, also über ein volles Intervall, integriert und die so entstehenden Teilintegrale zusammenfügt. Es ist daher wesentlich, daß die Summe der über die einzelnen K genommenen Integrale unmerklich bleibt; folglich ist h so klein zu wählen, daß die Differenzen längs jeder Zeile genügend rasch abnehmen. Hieraus fließt die Regel, daß man im Zweifelsfalle das Intervall lieber kleiner als größer zu wählen hat.

Im vorliegenden Falle soll, was sich als ausreichend erweisen wird, für h der zehnte Teil des ganzen Integrationsgebiets genommen werden; man gelangt damit zugleich, was immer angenehm ist, zu runden Werten für die Tafelargumente. Die untere Grenze des gesuchten Integrals ist, je nachdem man die Formel für $J(a + \frac{1}{2}h)_1$ oder $J(a)_1$ benutzt, von der Gestalt $a + \frac{1}{2}h$ oder a. Demgemäß ist, wenn zunächst nach $J(a + \frac{1}{2}h)_1$ gerechnet wird, zu setzen

$$h = 10000, \qquad a + \tfrac{1}{2}h = 100000, \qquad a = 95000.$$

Die nachstehende Tabelle enthält die wesentlichen Stücke der Rechnung. Unter A sind zur größeren Bequemlichkeit als Zeilennummern die Werte der Größe $A = a : 1000$ angesetzt, sodaß die auszuführende Integration von der Zeile 100 bis 200 auszudehnen ist. Daneben stehen unter $h(a, 0)$, I, II die nach (22) berechneten Werte von $hf(x)$ und deren Differenzen. Die in ( ) eingeschlossenen Zahlen sollen nachher besprochen werden. Die Rechnung ist mit einer fünfstelligen Logarithmentafel durchgeführt worden. Zur Erhöhung der Genauigkeit wurde dabei in der Zwischenrechnung die von den Proportionalteilen herrührende sechste Decimale mitgeführt und erst in den Schlußgrößen abgeworfen.

A	$h(a, 0)$	I	II
95	+ 872.47		
100	(+ 31)	− 756	
105	+ 864.91		+ 81
110		− 675	
115	+ 858.16		+ 67
120		− 608	
125	+ 852.08		+ 52
130		− 556	
135	+ 846.52		+ 48
140		− 508	
145	+ 841.44		+ 38
150		− 470	
155	+ 836.74		+ 35
160		− 435	
165	+ 832.39		+ 28
170		− 407	
175	+ 828.32		+ 29
180		− 378	
185	+ 824.54		+ 20
190		− 358	
195	+ 820.96		+ 22
200	(+ 8406.37)	− 336	
205	+ 817.60		

Aus dem Verhalten der zweiten Differenzen ist zu entnehmen, daß die dritten Differenzen für die Integration belanglos sind. Demnach nimmt die Verbesserung für die Zeile 100 den Wert

$$\left(+\frac{1}{24}\right) \cdot (-7.56) = -0.31$$

an. Der umgekehrte Betrag liefert das Summenglied der Zeile und ist zur Unterscheidung eingeklammert und der bequemeren Summation halber gleich in die Spalte $h(a, 0)$ eingetragen worden. Wird dann von Zeile 100 bis 200 summiert, so entsteht auf der Zeile 200 das ebenfalls in Klammern gesetzte Summenglied, welches mit der entsprechenden Verbesserung versehen den gesuchten Integralwert in der Gestalt

$$8406.37 + \left(+\frac{1}{24}\right) \cdot (-3.36) = 8406.37 - 0.14 = 8406.23$$

liefert.

§ 58. Soll die Rechnung nach $J(a)_1$ erfolgen, so ist zu setzen

$$h = 10000, \qquad a = 90000.$$

Die hierzu gehörigen und in derselben Weise wie vorhin berechneten Tafelgrößen sind nachstehend (ohne Klammern) zusammengestellt. Die Verbesserung für die Zeile 100 ist

$$-\frac{1}{12} \cdot \frac{-8.02 - 7.14}{2} = +0.631\,.$$

Die umgekehrte Verbesserung, mit dem halben Funktionswerte der

Zeile, nämlich + 434.295, additiv und subtraktiv verbunden, liefert die beiden in Klammern gesetzten Summenglieder zu den Zeilen 95 und 105. Daraus entstehen, wenn man aufsummiert, die beiden auf den Zeilen 195 und 205 stehenden Summenglieder.

A	$h\,f(a,0)$	I	II
90	+ 876.61		
95	(− 434.93)	− 802	
100	+ 868.59		+ 88
105	(+ 433.66)	− 714	
110	+ 861.45		+ 72
115		− 642	
120	+ 855.03		+ 63
125		− 579	
130	+ 849.24		+ 48
135		− 531	
140	+ 843.93		+ 43
145		− 488	
150	+ 839.05		+ 34
155		− 454	
160	+ 834.51		+ 35
165		− 419	
170	+ 830.32		+ 27
175		− 392	
180	+ 826.40		+ 25
185		− 367	
190	+ 822.73		+ 20
195	(+ 7996.32)	− 347	
200	+ 819.26		+ 22
205	(+ 8815.58)	− 325	
210	+ 816.01		

Werden die Summenglieder der Zeilen 195 und 205 nach Vorschrift verbunden und verbessert, so erhält man den gesuchten Integralwert

$$\frac{7996.32 + 8815.58}{2} - \frac{1}{12} \cdot \frac{-3.47 - 3.25}{2} = 8406.23\,.$$

Beide Rechnungen stimmen zufällig unter sich und mit dem anderweit bekannten korrekten Werte überein.

Das gewählte Intervall würde auch noch ausreichen, wenn man mit zwei Stellen mehr rechnen wollte, nur daß dann auch die dritten Differenzen in Betracht kommen. In diesem Falle hätte man, um die nötigen Differenzen dritter Ordnung zu erhalten, am Anfang und am Schluß der Tafel noch je einen Funktionswert hinzuzufügen. Ferner sieht man, daß die Vorschrift für $J(a + \frac{1}{2}h)_1$ etwas bequemer ist als die für $J(a)_1$, weil man bei jener einen Funktionswert spart und keine Mittel zu bilden hat.

§ 59. Die hier beschriebene Summenmethode ist den Rechnern in der Hauptsache zunächst durch die Mitteilungen von *Encke* geläufig

geworden, der das Verfahren im Jahre 1812 bei *Gauß* kennen lernte. Die Methode hat seither bei der Bearbeitung von Kometen und kleinen Planeten eine ausgiebige Verwendung gefunden: ohne ein derartiges Hilfsmittel wäre zur Zeit die Astronomie gezwungen bei der Mehrzahl der genannten Himmelskörper auf eine genauere Verfolgung ihrer Bewegungen einfach zu verzichten.

Es sind nun noch einige Umstände zu erörtern, die man bei dem Gebrauche der Summenmethode zu beachten hat. Der schwache Punkt des Verfahrens, das ja im übrigen an Geschmeidigkeit nichts zu wünschen übrig läßt, ist offenbar durch die Wirkung der Abrundungsfehler gegeben, die den Funktionswerten anhaften. Denn wenn auch im allgemeinen eine ziemlich ausgiebige Ausgleichung zwischen diesen teils positiven, teils negativen Fehlern stattfindet, so ist es doch unvermeidlich, daß bei gegebenem Intervall und gegebener Decimalenzahl die Unsicherheit der Integralwerte zunimmt, je weiter man in der Integraltafel fortschreitet. Man hat demnach, wenn die Ausdehnung des Integrationsgebietes vorgeschrieben ist, nach Festsetzung von h die Anzahl der mitzunehmenden Decimalen nach Möglichkeit derart zu wählen, daß die Wirkung der Abrundung für den Zweck der Rechnung unschädlich wird.

Ein anderer Punkt, der zu berücksichtigen ist, beruht darauf, daß für die Funktion $f(x)$ die *Taylor*sche Entwicklung nicht blos zwischen den Grenzen des gesuchten Integrals, sondern auch noch eine gewisse Strecke darüber hinaus zulässig sein muß, da ja hierauf die Zulässigkeit der zu Grunde gelegten Interpolationsformel beruht. So würde es z. B. nicht ohne weiteres angehen, das zwischen den Grenzen ± 1 genommene Integral

$$\int dx \sqrt{(1 - x^6)}$$

nach der Summenmethode auszuwerten, zumal die zu integrierende Funktion außerhalb der Integralgrenzen imaginär wird. Soll trotzdem die Summenmethode gebraucht werden, so müßte man das genannte Hindernis erst beseitigen, was in Fällen der vorliegenden Art häufig durch eine geeignete Transformation zu ermöglichen ist. Setzt man nämlich in dem angezogenen Beispiel $x = \sin v$ und schreibt

$$1 - x^6 = (1 - x^2)(1 + x^2 + x^4),$$

so nimmt das Integral die Gestalt

$$\int dv \cos v^2 \sqrt{(1 + \sin v^2 + \sin v^4)}$$

an, bei der offenbar jede Beschränkung bezüglich der Integrationsgrenzen fortgefallen ist.

Das vorstehende Integral gibt noch zu einer anderen Bemerkung Anlaß, die wir sogleich etwas allgemeiner fassen wollen. Es sei $f(x)$ eine stetige periodische Funktion mit der Periode 2π, sodaß

$$f(x + 2\pi) = f(x)$$

ist. Man teile die Peripherie 2π in n gleiche Teile, setze das Intervall h gleich dem Teilbogen $2\pi : n$, tabuliere die Funktion von der beliebig gewählten Stelle $x = a$ aus und stelle endlich die Tafel für das Integral J_1 auf. Will man nun den Wert des Integrals für eine Strecke kennen, die n Intervalle oder eine volle Periode umfaßt, so hat man die Werte von J_1 für die erste und letzte Zeile der Strecke aufzusuchen und voneinander zu subtrahieren. Hierbei fallen die von den Differenzen abhängenden Verbesserungen der Summengrößen in aller Strenge heraus, da sich, wie leicht zu sehen, die Werte der Differenzen nach n Intervallen periodisch wiederholen. Dadurch könnte man zu dem Schlusse verführt werden, daß das über eine volle Periode erstreckte Integral von $f(x)$ in aller Strenge gefunden werde, wenn man für n äquidistante Punkte der Peripherie die Werte von $f(x)$ rechnet und ihre Summe mit h multipliziert oder mit anderen Worten, wenn man das arithmetische Mittel dieser Funktionswerte mit 2π multipliziert. Daß dieser Schluß falsch ist, läßt sich leicht an Beispielen zeigen; überdies würde er besagen, daß das genannte arithmetische Mittel von der Teilungszahl n unabhängig sei. Der Widerspruch rührt daher, daß sich zwar die von den benutzten Differenzen abhängenden Verbesserungen der Summengrößen, aber nicht die aus den vernachlässigten Korrektionen K entspringenden Bestandteile herausheben.

Die vorstehende Bemerkung enthält offenbar eine Verstärkung der Gründe, welche die Wahl kleiner Tafelintervalle ratsam machen.

§ 60. Die bisherige Untersuchung bezog sich ausschließlich auf die Ausführung von Quadraturen. Wichtiger noch als diese ist aber die Anwendung der Summenmethode auf die Integration von Differentialgleichungen. Da der hierbei einzuschlagende Weg wesentlich von der Beschaffenheit der vorgelegten Differentialgleichungen abhängt, und da man ferner sich die Besonderheiten dieser Gleichungen nach Möglichkeit zu nutze machen wird, so beschränke ich mich darauf, den Grundgedanken des Verfahrens an einem Beispiel auseinander zu setzen.

Die Größe y sei als Funktion von x durch die Differentialgleichung

$$\frac{d^2y}{dx^2}\left(x\frac{dy}{dx} - y\right) = 1$$

definiert, mit der Bedingung, daß für $x = 1$

$$y = 1, \quad \frac{dy}{dx} = 0$$

werden solle. Ferner werde verlangt, daß man den numerischen Verlauf von y durch eine Tafel darstelle, die von $x = 1$ ausgehend mit einem passenden Intervall fortschreitet. Um die Rechnung etwas bequemer zu gestalten setzen wir

$$x = 1 + th, \quad y = 1 + z, \quad \frac{dz}{dt} = z_1, \quad \frac{dz_1}{dt} = z_2,$$

sodaß die Differentialgleichung die Gestalt

$$z_2 [z_1 + h(tz_1 - z - 1)] = h^3 \tag{23}$$

annimmt. Die verlangte Tafel ist dann für ganzzahlige t von $t = 0$ an zu rechnen, ferner gelten für $t = 0$ die Anfangswerte $z = 0$ und $z_1 = 0$. Setzt man nun noch zur Abkürzung

$$T = h^3 : [z_1 + h(tz_1 - z - 1)], \tag{24}$$

so wird $z_2 = T$. Daraus folgt

$$z_1 = \int T\, dt, \qquad z = \int dt \int T\, dt, \tag{25}$$

wo die Integrationen jedesmal mit $t = 0$ zu beginnen haben.

Dies vorausgeschickt nehmen wir jetzt an, es seien bereits für z und z_1 genäherte Wertreihen vorhanden. Dann kann man damit nach (24) die Werte von T für die vorgeschriebenen Tafelargumente berechnen und die in (25) geforderten Integrationen numerisch durchführen. Hierdurch entstehen für z und z_1 neue Wertreihen, die der Wahrheit näher kommen werden, als die ursprünglichen, falls das Intervall h hinreichend klein gewählt worden war, denn es gehen dann die Fehler der ursprünglichen Werte von z und z_1 verkleinert in das Ergebnis der Integrationen ein. Mit den aus (25) gefundenen verbesserten Werten wird nun die ganze Rechnung wiederholt und dieses Hinundher solange fortgesetzt, bis die Rechnung zum Stehen kommt, bis also die Durchrechnung von (24) und (25) keine weitere Verbesserung mehr erzeugt.

Um die für den Beginn der Rechnung nötige Näherung anzusetzen beschränken wir uns für den Augenblick auf die nächste Umgebung der Stelle $t = 0$ und verzichten dabei absichtlich auf die Hilfen, die man aus der besonderen Beschaffenheit der vorgelegten Differentialgleichung ziehen könnte. Da in diesem Falle die Änderungen von t, z, z_1 und entsprechend die Änderungen von T klein sind, so setzen wir in T zunächst einfach die Anfangswerte $t = 0$, $z = 0$, $z_1 = 0$ unter Vernachlässigung ihrer Änderungen ein und erhalten damit $T = -h^2$. Daraus folgt nach (25)

$$z_1 = -h^2 t, \qquad z = -\tfrac{1}{2} h^2 t^2. \tag{26}$$

Mit dieser Näherung, die sich ohne weitere Umstände ergibt, beginnen

wir die eigentliche Rechnung, für welche noch die zu benutzende Stellenzahl und das Intervall h festzusetzen ist. Die Stellenzahl für z wird, wenn es sich nicht wie hier um ein bloßes Paradigma handelt, im allgemeinen durch den Zweck vorgeschrieben sein, dem die Integration der vorgelegten Differentialgleichung dienen soll. Dagegen lassen sich für die Bemessung von h keine allgemeine Regeln im voraus aufstellen. Nimmt man h sehr klein, so verlaufen die einzelnen Operationen sehr bequem, weil dann die Verbesserung der Summengrößen nur ein oder vielleicht zwei Glieder umfaßt. Dafür schreitet aber die Herstellung der verlangten Tafel nur langsam fort, ferner wird die Arbeit zur Berechnung der Größen T in unliebsamer Weise namentlich dann vermehrt, wenn T ein verwickelter oder weitläufiger Ausdruck ist. Auf der anderen Seite zwingt ein großes h dazu, in der Berücksichtigung der Differenzenreihen weiter zu gehen, als angenehm ist, und gefährdet überdies die Sicherheit der Rechnung. Man muß daher, wenn für die Wahl von h keine Anhaltspunkte vorliegen, das Intervall zunächst auf gut Glück ansetzen und die Rechnung wenn auch nur roh beginnen. Stellt sich hierbei heraus, daß h zu groß oder zu klein genommen war, so wird man mit einem geeigneteren h nochmals anfangen. Außerdem kann auch der Fall eintreten, daß das anfangs eingeführte Intervall in einem späteren Teile der zu berechnenden Tafel unzweckmäßig wird und abgeändert werden muß. Ist nun das Intervall gewählt, so kann man sich auch schlüssig machen, wie viele Stellen bei der Ansetzung von T nötig sind: hierbei kommt wegen der Wirkung der Abrundungsfehler in den Summenreihen außer der Stellenzahl von z auch noch der Wert des Arguments t in Betracht, bis zu welchem die verlangte Tafel ausgedehnt werden soll.

Im vorliegenden Falle handelt es sich nur um ein Paradigma, deshalb soll $h = 0.1$ und T fünfstellig genommen werden. Demgemäß sind weiterhin die von t abhängenden Größen immer in Einheiten der fünften Decimale angesetzt.

§ 61. Bei Beginn der Rechnung gehen wir zunächst nur bis $t = 2$, nehmen aber wegen der zu bildenden Differenzen noch die beiden Intervalle vor $t = 0$ mit, sodaß die Rechnung mit den beiden nach (26) gebildeten Wertreihen für z und z_1, nämlich

$t =$	-2	-1	0	$+1$	$+2$
$z =$	-2000	-500	0	-500	-2000
$z_1 =$	$+2000$	$+1000$	0	-1000	-2000

anfängt. Hiermit sind nach (24) unter Benutzung vierstelliger Logarithmen die Werte von T berechnet, die man in der nachstehenden, noch näher zu erläuternden, Tabelle A unter der Überschrift (0) angesetzt findet.

Tabelle A.

t	(— 2)	(— 1)	(0)	(1)	(2)	z	z_1
		+ 2833		(+ 125)			
— 2	— 2038		— 1220		(— 10)	— 2140	+ 2213
		+ 1613		+ 115			
— 1	— 425		— 1105		— 10	— 517	+ 1051
		+ 508		+ 105			
0	+ 83		— 1000		— 10	0	0
		— 492		+ 95			
+ 1	— 409		— 905		— 10	— 484	— 952
		— 1397		+ 85			
+ 2	— 1806		— 820		(— 10)	— 1874	— 1814
		— 2217		(+ 75)			

Linkerhand sind der Bequemlichkeit halber die Argumente t hinzugefügt, die jetzt zugleich als Zeilennummern dienen. Rechts von der Spalte (0) stehen die Differenzen bis zur zweiten Ordnung, wobei die eingeklammerten Zahlen extrapoliert worden sind, und zwar unter der hier von selbst gegebenen Annahme, daß die zweiten Differenzen konstant seien. Die zur Bildung der beiden Integrale z und z_1 erforderlichen Summengrößen stehen unter (— 1) und (— 2) und beruhen auf der Bedingung, daß die Integrale auf der Zeile $t = 0$ verschwinden sollen. Aus den Summengrößen ergeben sich die unter z und z_1 angesetzten Integralwerte, mit denen nunmehr die zweite Näherung durchzurechnen ist. Das Ergebnis ist in der nachfolgenden Tabelle B enthalten, die ebenso zu verstehen ist, wie die Tabelle A, nur daß jetzt beim Extrapolieren die dritten Differenzen in der ersichtlich gemachten Weise berücksichtigt worden sind.

Tabelle B.

t	(— 2)	(— 1)	(0)	(1)	(2)	(3)	z	z_1
		+ 2866		(+ 166)				
— 2	— 2044		— 1247		(— 30)		— 2148	+ 2230
		+ 1619		+ 136		(+ 5)		
— 1	— 425		— 1111		— 25		— 518	+ 1053
		+ 508		+ 111		+ 5		
0	+ 83		— 1000		— 20		0	0
		— 492		+ 91		+ 4		
+ 1	— 409		— 909		— 16		— 485	— 953
		— 1401		+ 75		(+ 4)		
+ 2	— 1810		— 834		(— 12)		— 1879	— 1824
		— 2235		(+ 63)				

Vergleicht man die drei Wertsysteme, die bisher für z und z_1 vorgekommen sind, so erkennt man, daß der Näherungsprozeß rasch konvergiert, und daß die nochmalige Durchrechnung die endgültigen Werte von T liefern wird. Demgemäß legt man jetzt das Blatt für die endgültige Rechnung an, wie aus der Tabelle C zu ersehen ist.

Tabelle C.

t	(— 2)	(— 1)	(0)	(1)	(2)	(3)	z	z_1
— 2	—		— 1250					
		....		+ 139				
— 1	-		— 1111		— 28			
		+ 508		+ 111		+ 8		
0	+ 83		— 1000		— 20		0	0
		— 492		+ 91		+ 5		
+ 1	— 409		— 909		— 15		— 485	— 953
		— 1401		+ 76		(+ 6)		
+ 2	— 1810		— 833		(— 9)		— 1879	— 1823
		— 2234		(+ 67)		(+ 6)		
+ 3	— 4044		(— 766)		(— 3)		(— 4108)	(— 2622)
		(— 3000)		(+ 64)				

Zunächst trägt man die Werte von T ein, die sich mit den zuletzt gefundenen Werten von z und z_1 ergeben. Dazu bildet man die Differenzen und Summen, soweit sie sich bilden lassen und für die mit $t = 0$ beginnende Tafel der gesuchten Funktion z nötig sind. Weiter extrapoliert man die in ( ) eingeschlossenen Glieder des Differenzenschemas, wobei in Ermangelung anderer Anhaltspunkte die dritten Differenzen konstant und gleich dem Mittel der bereits vorhandenen dritten Differenzen gesetzt worden sind. Daraus erhält man die Integrale z und z_1, und zwar bis $t = 2$ mit endgültigen, für $t = 3$ hingegen mit vorläufigen Werten, die zur Unterscheidung in ( ) gefaßt sind. Berechnet man hiermit T für die Zeile $t = 3$, so erhält man statt des extrapolierten Betrages — 766 die Zahl — 769, die als endgültig betrachtet werden darf, da vorherzusehen ist, daß eine nochmalige Durchrechnung der Zeile wegen der Kleinheit der vorhandenen Abweichung an dem gefundenen T nichts ändern würde. In der Tat erhält man jetzt

$$z = -4108, \quad z_1 = -2623.$$

Um das Rechenschema nicht mit Ziffern zu belasten, die nach Ermittelung der endgültigen Zahlen kein weiteres Interesse mehr besitzen, kann man die Rechnung für die vorläufigen Größen auf besonderen Zetteln ausschreiben, sodaß in das Schema nur die endgültigen Größen eingetragen werden. Oder aber man trägt die vorläufigen Werte wirklich ein, aber mit Bleistift, und übermalt sie nachher mit gutschwarzer Tinte; die Bleistiftzüge werden später mit weichem Gummi weggewischt.

Es bedarf keiner besonderen Auseinandersetzung, wie der Prozeß weiter läuft, sobald man erst einmal die Rechnung in Gang gebracht hat: ist eine Zeile, wie z. B. vorhin die Zeile $t = 3$, fertig gerechnet, so extrapoliert man für die nächste Zeile, ermittelt daraus vorläufige Werte für die Integrale z, z_1 und erhält im allgemeinen mit einmaliger Durchrechnung die endgültigen Zahlen der Zeile, weil schon der Fehler der Extrapolation stark verkleinert in die vorläufigen Werte von

z und z_1 eingeht. Infolgedessen nimmt das ganze Verfahren, sobald erst die Vorbereitungen erledigt sind, den Charakter einer direkten Rechnung an.

§ 26. Bei der vorstehenden Beschreibung des Rechnungsganges war absichtlich die Erleichterung beiseite gelassen worden, die man häufig zur Abkürzung der Vorbereitung aus der besonderen Beschaffenheit der vorgelegten Differentialgleichung ziehen kann. Setzt man nun

$$U = z_1 + h(tz_1 - z - 1) \tag{27}$$

und bezeichnet die Ableitungen von U ebenfalls durch Indices, so nimmt die Differentialgleichung (23) die Gestalt

$$z_2 U = h^3 \tag{28}$$

an. Differentiiert man (27) und (28), so entsteht

$$U_1 = z_2 + thz_2,$$
$$U_2 = z_3 + thz_3 + hz_2,$$
u. s. w.

$$0 = z_3 U + z_2 U_1,$$
$$0 = z_4 U + 2 z_3 U_1 + z_2 U_2,$$
u. s. w.

Daraus erhält man für $t = 0$ die Werte

$$z = z_1 = 0, \quad z_2 = -h^2, \quad z_3 = h^3, \quad z_4 = -2h^4, \cdots,$$

sodaß man die *Taylor*schen Reihen

$$z = -\tfrac{1}{2}(th)^2 + \tfrac{1}{6}(th)^3 - \tfrac{1}{12}(th)^4 + \cdots,$$
$$z_1 = -h(th) + \tfrac{1}{2}h(th)^2 - \tfrac{1}{3}h(th)^3 + \cdots$$

bilden kann. Berücksichtigt man nur die hingeschriebenen Glieder, so kann man für den Beginn der Rechnung sogleich das Wertsystem

$t =$	-2	-1	0	$+1$	$+2$
$z =$	-2146	-518	0	-484	-1880
$z_1 =$	$+2227$	$+1053$	0	-053	-1827

ansetzen, das der Wahrheit bereits sehr nahe kommt.

Die Heranziehung der *Taylor*schen Reihe für den Beginn der Rechnung setzt natürlich voraus, daß sich die nötigen Ableitungen ohne besondere Mühe ermitteln lassen, da es andernfalls bequemer ist, nach dem früheren Verfahren einige Näherungen mehr durchzurechnen.

Da das benutzte Beispiel sich ohne Schwierigkeit analytisch integrieren läßt, so möge hier noch die Vergleichung der numerischen Integration mit dem analytischen Ausdrucke folgen. Eliminiert man

aus den oben aufgestellten Gleichungen

$$z_2 U = h^3, \quad U_1 = z_2(1 + th), \quad 0 = z_3 U + z_2 U_1$$

die Größen U und U_1, so wird

$$0 = z_3 h^3 + z_2^3 (1 + th)$$

oder

$$-h^4 z_3 : z_2^3 = h(1 + th),$$

woraus durch Integration

$$h^4 : z_2^2 = (1 + th)^2$$

folgt, da die hinzuzufügende willkürliche Integrationskonstante verschwindet, weil für $t = 0$ nach (27) und (28) $U = -h$ und $z_2 = -h^2$ wird. Durch Wurzelausziehung entsteht die Gleichung

$$z_2 = -h^2 : (1 + th),$$

deren Integration unter Berücksichtigung der Anfangsbedingungen erst

$$z_1 = -h \log(1 + th)$$

und weiter

$$z = th - (1 + th) \log(1 + th)$$

liefert. Bezeichnet man die hieraus folgenden strengen Werte von z und z_1 mit Z und Z_1, so entsteht durch Vergleichung mit den rein numerisch gefundenen Zahlen das nachstehende Täfelchen:

$t =$	-2	-1	0	$+1$	$+2$	$+3$
$Z =$	-2148.5	-517.6	0	-484.1	-1878.6	-4107.4
$z =$	-2148	-518	0	-485	-1879	-4108
$Z_1 =$	$+2231.4$	$+1053.6$	0	-953.1	-1823.2	-2623.6
$z_1 =$	$+2230$	$+1053$	0	-953	-1823	-2623

Die Unterschiede sind offenbar nicht größer, als man sie nach der angewandten Rechnungsschärfe zn erwarten hat.

§ 63. Ist für ein bestimmtes Integral oder für die Lösung einer Differentialgleichung eine zur numerischen Rechnung taugliche analytische Darstellung, sei es durch geschlossene Ausdrücke, sei es durch Reihen oder dergleichen, erhalten worden, so gilt es gemeinhin als selbstverständlich, daß jene Darstellung bei der Rechnung auch tatsächlich zu Grunde gelegt werde. Daß ein solches Verfahren im allgemeinen zweckmäßig ist steht außer Zweifel, es ist jedoch nicht überflüssig darauf hinzuweisen, daß auch Ausnahmen vorkommen können und daß im besondern die rein numerischen Methoden, die gewöhnlich nur als Notbehelf gelten, unter Umständen rascher und glatter arbeiten als die Rechnung nach den analytischen Ausdrücken.

Einfache Beispiele werden genügen, um die Berechtigung dieser Bemerkung zu erhärten.

Angenommen, es handele sich darum, eine achtstellige Tafel von $\operatorname{Log} x$ für die ganzzahligen x von 10000 bis 99900 völlig neu zu berechnen. Dann ist, wenn M den Modulus der gemeinen Logarithmen bedeutet,

$$\operatorname{Log} x = \int \frac{M}{x}\, dx.$$

Tabuliert man nun den Quotienten $M : x$ mit dem Intervall $h = 1$ auf 20 Stellen und führt dann, wie es oben gelehrt wurde, die numerische Integration aus und zwar einmal von $x = 10000$ aus vorwärts bis 55000, sodann von $x = 100000$ aus rückwärts bis 55000, so kann die Anhäufung der Abrundungsfehler, selbst wenn alle Fehler in derselben Richtung liegen, nicht über 22500 Einheiten der letzten Stelle hinausgehen, d. h. in dem denkbar ungünstigten, tatsächlich aber nicht eintretenden Falle würde die 16te Decimale um rund zwei Einheiten fehlerhaft ausfallen können. Hierbei umfaßt, falls nach der Formel für $J(a + \frac{1}{2}h)_1$ gerechnet wird, der für den einzelnen Logarithmus nötige Arbeitsaufwand folgende Bestandteile: 1) Berechnung des Quotienten $M : x$, welcher der Zeile des Logarithmus unmittelbar vorhergeht; 2) Bildung des Summengliedes und der Differenzen, welche auf der Zeile des Logarithmus oder auf der unmittelbar vorhergehenden Hauptzeile stehen; 3) Verbesserung des Summengliedes. Benutzt man die Formel für $J(a)_1$ und berechnet sogleich die Quotienten $M : 2x$, so treten als Mehrarbeit noch die zur Bildung der Zwischengrößen nötigen Summationen hinzu. Beachtet man nun, daß die geforderten Operationen sich mit den heutigen Rechenmaschinen rasch und glatt erledigen lassen, und vergegenwärtigt man sich den Zeitaufwand, der bei den gewöhnlich für die Tabulierung von $\operatorname{Log} x$ empfohlenen Vorschriften nötig sein würde, so kommt man zu dem Ergebnis, daß im vorliegenden Falle das an sich höchst kunstlose Summenverfahren die anderen Methoden an Zweckmäßigkeit hinter sich läßt.

Ein anderes Beispiel ist die in der Fehlertheorie auftretende Transcendente

$$y = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-t^2)\, dt.$$

Die Entwickelung der Exponentialgröße nebst darauf folgender Integration führt auf die beständig konvergente Reihe

$$y = \frac{2}{\sqrt{\pi}} \left[x - \frac{x^3}{1!\,3} + \frac{x^5}{2!\,5} - \cdots \right];$$

außerdem ist, wie hier nicht nachgewiesen werden soll, eine für größere

Werte von x brauchbare semikonvergente Entwickelung vorhanden. Auch hier ist, sobald y mit einer größeren Stellenzahl tabuliert werden soll, die numerische Integration rationeller, als die Benutzung der Reihen. Es ist nicht nötig dies näher zu begründen, da *Th. von Oppolzer* in dem oben (§ 51) genannten Lehrbuche die in Rede stehende Rechnung tatsächlich ausgeführt hat.

Die vorstehenden Beispiele lassen erkennen, daß es sich bei der Tabulierung von Integralen häufig der Mühe lohnen wird zu überlegen, ob nicht etwa der rein numerische Weg rascher zum Ziele führt als die Rechnung nach analytischen Ausdrücken.

Fünfter Abschnitt.

Numerische Integration: Viereckverbesserung.

§ 64. Bezeichnet man, da es sich im folgenden immer nur um einfache Integrale handelt, das Integral der vorgelegten Funktion kurz mit $J(x)$, so ist $J(x)$ zugleich der Flächeninhalt der Figur, die begrenzt wird von der Abscissenachse, von der Kurve $y = f(x)$, von der Ordinate y und von einer gewissen Anfangsordinate. Weiterhin soll, wenn von dem Flächeninhalt einer Kurve die Rede ist, darunter immer der Inhalt einer solchen Figur verstanden werden. Läßt man x um das Intervall h wachsen, so stellt die Differenz

$$J(x+h) - J(x)$$

den Inhalt des „*Flächenstreifens*" dar, der seitlich von den zu x und $x+h$ gehörigen Ordinaten begrenzt wird.

Bei der Herleitung der in diesem Abschnitte zu behandelnden Formeln soll zunächst an die Ergebnisse der Summenmethode angeknüpft werden. Die Gleichung (7) aus § 54 läßt sich, da der Anfangskoeffizient, wie wir wissen, gleich Eins ist, in der Gestalt

$$J(a) = h\,(a, -1) + \sum_q c_q h\,(a, 2q-1), \qquad (q = 1, 2, \cdots)$$

schreiben. Verschiebt man a um ein Intervall und bildet die erste Differenz, so entsteht

$$J(a+h) - J(a) = h\,(a + \tfrac{1}{2}, 0) + \sum_q c_q h\,(a + \tfrac{1}{2}, 2q). \qquad (1)$$

Die linke Seite ist der Inhalt des Flächenstreifens, der zwischen den Abscissen a und $a+h$ enthalten ist. Zieht man nun zu dem Kurvenbogen, der den Streifen abschließt, die Sehne, so wird dadurch ein Paralleltrapez T und ein Segment S erzeugt. Die parallelen Seiten des Trapezes sind durch die beiden Ordinaten $(a, 0)$ und $(a+1, 0)$ gegeben, während die Breite des Trapezes gleich dem Intervall h ist. Folglich erhält man für den Trapezinhalt den Ausdruck

$$T = \tfrac{1}{2}h\,[(a, 0) + (a+1, 0)] = h\,(a + \tfrac{1}{2}, 0).$$

Dies ist aber nichts anderes als das Anfangsglied der rechten Seite

von (1), woraus weiter folgt, daß der Inhalt des Segments durch die in (1) vorkommende Reihe

$$S = \sum_q e_q(a + \tfrac{1}{2}, 2q)$$

gegeben ist.

In ähnlicher Weise läßt sich die Formel (4) in § 53 deuten, indem man zunächst

$$J(a + \tfrac{1}{2}h) = h(a + \tfrac{1}{2}, -1) + \sum_q b_q h(a + \tfrac{1}{2}, 2q - 1)$$

schreibt. Die erste Differenz hierzu liefert

$$J(a + \tfrac{1}{2}h) - J(a - \tfrac{1}{2}h) = h(a, 0) + \sum_q b_q h(a, 2q). \qquad (2)$$

Die linke Seite ist der Flächenstreifen zwischen den Abscissen $x = a \pm \frac{1}{2}h$. Konstruiert man die in der Mitte des Streifens liegende, zur Abscisse $x = a$ gehörende *„Mittelordinate"* und zieht durch ihren Endpunkt eine Parallele zur Abscissenachse, so entsteht ein Rechteck R mit dem Inhalt $h(a, 0)$, das dem Anfangsglied auf der rechten Seite von (2) entspricht. Bezeichnet man ferner die Differenz „Flächenstreifen minus Rechteck" als den *„Exceß E"*, so ist offenbar

$$E = \sum_q b_q h(a, 2q).$$

In den beiden Formeln (1) und (2) ist der Inhalt von Segment oder Exceß, der als Verbesserung zu dem als erste Näherung dienenden Viereck — Trapez oder Rechteck — hinzuzufügen ist, von Ordinaten abhängig gemacht, die zum Teil *außerhalb* des Integrationsgebietes liegen, da ja die in (1) und (2) auftretenden Differenzen auf mehr als einem Intervall beruhen. Diese Verbesserungen sollen nun so umgeformt werden, daß sie nur von dem Verlaufe der Kurve *innerhalb* des Integrationsgebietes abhängen.

§ 65. Die Funktion $f(x) = f(a + th)$ sei nach dem *Taylor*schen Satze in die Reihe

$$f(x) = f(a + th) = A_0 + A_1 th + A_2 (th)^2 + \cdots \qquad (3)$$

entwickelt. Ferner werde die q^{te} Ableitung von $f(x)$ mit $f(x)_q$ bezeichnet und die Integrationskonstante in $J(x)$ so angenommen, daß $J(a)$ verschwindet. Dann lassen sich zunächst folgende Gleichungen ansetzen:

$$\begin{aligned}
J(a + th) &= A_0 th + \tfrac{1}{2} A_1 (th)^2 + \tfrac{1}{3} A_2 (th)^3 + \cdots,\\
f(a + th) &= A_0 + A_1 th + A_2 (th)^2 + \cdots,\\
f(a + th)_1 &= A_1 + 2 A_2 th + 3 A_3 (th)^2 + \cdots,\\
f(a + th)_3 &= 6 A_3 + 24 A_4 th + 60 A_5 (th)^2 + \cdots,
\end{aligned}$$

u. s. w.

Daraus entstehen, wenn man $t = \pm \frac{1}{2}$ setzt und zur Abkürzung

$$f(a + \tfrac{1}{2}h)_q = F_q, \qquad f(a - \tfrac{1}{2}h)_q = G_q$$

schreibt, Gleichungen von der Form

$$\left.\begin{aligned}
J(a + \tfrac{1}{2}h) - J(a - \tfrac{1}{2}h) &= A_0 h + b_1 A_2 h^3 + b_2 A_4 h^5 + \cdots,\\
\tfrac{1}{2}h(F_0 + G_0) &= A_0 h + c_1 A_2 h^3 + c_2 A_4 h^5 + \cdots,\\
h^2(F_1 - G_1) &= \qquad\quad d_1 A_2 h^3 + d_2 A_4 h^5 + \cdots,\\
h^4(F_3 - G_3) &= \qquad\qquad\qquad\quad e_2 A_4 h^5 + \cdots,\\
\text{u. s. w.}
\end{aligned}\right\} \quad (4)$$

wo die Größen $b, c, d, \cdots$ reine Zahlenkoeffizienten bedeuten, und von den Größen f_q außer $f(x)_0 = f(x)$ nur die Ableitungen ungerader Ordnung auftreten. Die vorstehenden Gleichungen sollen jetzt vermittelst der Multiplikatoren

$$1, \quad -1, \quad -B_1, \quad -B_2, \quad -B_3, \cdots$$

verbunden und diese so bestimmt werden, daß auf der rechten Seite die Größen $A_0, A_2, \cdots$ herausfallen. Die Multiplikatoren B sind dann reine Zahlengrößen und man erhält

$$J(a+\tfrac{1}{2}h) - J(a-\tfrac{1}{2}h) = \tfrac{1}{2}h(F_0 + G_0) + \sum_q B_q h^{2q}(F_{2q-1} - G_{2q-1}), \quad (q = 1, 2, \cdots). \quad (5)$$

Geht man bei der Elimination der A-Größen nur bis A_{2q}, so reicht die rechte Seite von (5) zunächst bis zu dem Gliede

$$B_q h^{2q}(F_{2q-1} - G_{2q-1}),$$

daran schließt sich dann ein Rest, der als eine Potenzreihe von h erscheint, deren niedrigstes Glied nach h von der Ordnung $2q + 3$ ist.

Um die Koeffizienten B zu bestimmen, werde für den Augenblick $f(x)$ durch die Exponentialfunktion $\exp x$ ersetzt und a gleich Null genommen. Dann wird mit den Abkürzungen

$$2C = \exp(\tfrac{1}{2}h) + \exp(-\tfrac{1}{2}h), \qquad D = \exp(\tfrac{1}{2}h) - \exp(-\tfrac{1}{2}h)$$

der Reihe nach

$$J(a + \tfrac{1}{2}h) - J(a - \tfrac{1}{2}h) = D,$$

$$F_0 + G_0 = 2C, \quad F_q - G_q = D.$$

Die Substitution dieser Ausdrücke in (5) liefert die Gleichung

$$hC : D = 1 - B_1 h^2 - B_2 h^4 - \cdots.$$

Entwickelt man hierin die linke Seite durch gewöhnliche Division, so erhält man

$$B_1 = -1 : 12, \qquad B_2 = 1 : 720, \qquad \text{u. s. w.}$$

Die vorstehenden Werte genügen, da man kaum jemals Veranlassung haben wird, über das Glied mit B_2 hinauszugehen.

§ 66. Die linke Seite von (5) ist der Inhalt des Flächenstreifens zwischen den Abscissen $a \pm \frac{1}{2}h$, das Anfangsglied rechts gibt den Inhalt des zugehörigen Trapezes an, sodaß der Segmentinhalt durch die Reihe mit den Koeffizienten B_1, B_2, $\cdot\cdot$ ausgedrückt ist. Ersetzt man a durch $a + \frac{1}{2}h$ und beachtet die Bedeutung der Abkürzungen T und S, so gelangt man zu den Gleichungen

$$\left.\begin{aligned} \text{Flächenstreifen} &= J(a+h) - J(a) = T + S, \\ \text{Trapez } T &= \tfrac{1}{2}h\,[f(a) + f(a+h)], \\ \text{Segment } S &= \sum_q B_q h^{2q}\,[f(a+h)_{2q-1} - f(a)_{2q-1}]\cdot \end{aligned}\right\} \qquad (6)$$

Sollen die vorstehenden Formeln für die numerische Rechnung hinlänglich bequem sein, so darf die von dem Segment herrührende Verbesserung nur kleine Beträge erreichen. Demnach wird man, wenn das Integral

$$J = \int_a^b f(x)\,dx$$

gesucht ist, das von $x = a$ bis $x = b$ reichende Integrationsgebiet gewöhnlich erst in n gleiche Strecken von der Länge $h = (b-a):n$ zerlegen, dann die so entstehenden Streifen nach (6) berechnen und summieren. Hierbei hat man es in der Hand, durch angemessene Wahl von n die Segmentwerte auf bequeme Beträge herunterzubringen. Bildet man demgemäß mit den zu den Endpunkten der Teilstrecken gehörigen Ordinaten

$$f(a), \qquad f(a+h), \quad \cdot\cdot \quad f(b-h), \qquad f(b)$$

den Ausdruck

$$U = [f(a) + f(a+h) + \cdot\cdot + f(b)] - \tfrac{1}{2}[f(a) + f(b)]$$

und bezeichnet mit (T) und (S) die Summen der Trapeze und Segmente, so wird, weil sich bei der Summation die Ableitungen der Zwischenordinaten in (6) herausheben,

$$\left.\begin{aligned} J &= (T) + (S), \\ (T) &= \frac{b-a}{n}\,U, \\ (S) &= -\frac{(b-a)^2}{12\,n^2}\,[f(b)_1 - f(a)_1] + \frac{(b-a)^4}{720\,n^4}\,[f(b)_3 - f(a)_3] - \cdot\cdot \end{aligned}\right\} \qquad (7)$$

Die Segmentsumme ist also nur von den Ableitungen der beiden äußersten Ordinaten abhängig.

Die Formel (7) geht auf *Euler* und *Mac Laurin* zurück und ist ursprünglich für einen entgegengesetzten Zweck abgeleitet worden, nämlich zur Summierung der in U auftretenden Reihe durch ein bestimmtes Integral.

§ 67. Um die Formel für die Rechteckverbesserung abzuleiten, gehen wir auf das Gleichungssystem (4) zurück und ersetzen darin die zweite Gleichung durch den Ausdruck für die Mittelordinate, so daß das System

$$
\begin{aligned}
J(a + \tfrac{1}{2}h) - J(a - \tfrac{1}{2}h) &= A_0 h + b_1 A_2 h^3 + b_2 A_4 h^5 + \cdots \\
hf(a) &= A_0 h \\
h^2(F_1 - G_1) &= \qquad d_1 A_2 h^3 + d_2 A_4 h^5 + \cdots, \\
&\text{u. s. w.}
\end{aligned}
$$

entsteht. Eliminiert man die A vermittelst der Multiplikatoren

$$1, \; -1, \; -C_1, \; -C_2, \cdots,$$

so entsteht die Gleichung

$$J(a + \tfrac{1}{2}h) - J(a - \tfrac{1}{2}h) = hf(a) + \sum_q C_q h^{2q}(F_{2q-1} - G_{2q-1}). \quad (8)$$

Die Substitution von $\exp x$ für $f(x)$ liefert ähnlich wie vorhin zur Bestimmung der C-Größen die Reihe

$$h : D = 1 - C_1 h^2 - C_2 h^4 - \cdots,$$

woraus

$$C_1 = 1 : 24, \qquad C_2 = -7 : 5760, \cdots$$

folgt.

Der Flächenstreifen wird jetzt durch das mit der Mittelordinate konstruierte Rechteck $R = hf(a)$ und den Exceß E dargestellt. Ersetzt man wieder a durch $a + \frac{1}{2}h$, so erhält man

$$
\left.
\begin{aligned}
\text{Flächenstreifen} &= J(a + h) - J(a) = R + E, \\
\text{Rechteck } R &= hf(a + \tfrac{1}{2}h), \\
\text{Exceß } E &= \sum C_q h^{2q} [f(a + h)_{2q-1} - f(a)_{2q-1}].
\end{aligned}
\right\} \quad (9)
$$

Um das Integral

$$J = \int_a^b f(x)\,dx$$

zu erhalten, teilt man die Strecke $b - a$ wieder in die n gleichen Intervalle $h = (b - a) : n$, berechnet aus den Mittelordinaten der einzelnen Streifen die Summe

$$V = f(a + \tfrac{1}{2}h) + f(a + \tfrac{3}{2}h) + \cdots + f(b - \tfrac{1}{2}h),$$

und erhält, wenn (R) und (E) die Summen der Rechtecke und Excesse bedeuten, die Darstellung

$$
\left.
\begin{aligned}
J &= (R) + (S), \\
(R) &= \frac{b - a}{n} V, \\
(E) &= \frac{(b - a)^2}{24 n^2}[f(b)_1 - f(a)_1] - \frac{7(b - a)^4}{5760 n^4}[f(b)_3 - f(a)_3] + \cdots
\end{aligned}
\right\} \quad (10)
$$

Die Verbesserung hängt wiederum nur von den Ableitungen der beiden äußersten Ordinaten ab.

§ 68. Die Formeln (7) und (10) sind offenbar zur numerischen Integration recht gut brauchbar, sobald man die Werte der vorkommenden Ableitungen ohne Weitläufigkeit berechnen kann. Man vermag in diesem Falle auch sogleich zu überschlagen, wie groß die Teilungszahl n zu nehmen ist, damit die Summe der Segmente oder Excesse von vornherein den Charakter einer kleinen und rasch zu berechnenden Verbesserung annehme. Die Vergleichung mit der Summenmethode läßt erkennen, daß die Hauptglieder bei dem einen wie dem anderen Verfahren dieselben sind, daß also der Unterschied nur in der Form der Verbesserungen besteht. Da die Summenmethode hierbei den Verlauf der Kurve außerhalb der Strecke $b-a$ heranzieht, so kann sie, wie bereits erwähnt wurde, unter Umständen versagen, indem irreguläre Stellen, die außerhalb der Strecke $b-a$ in der Nähe von a oder b liegen, die für die Verbesserung aufgestellten Reihen zur numerischen Rechnung unbrauchbar machen. Bei der Viereckverbesserung sind dagegen, wie aus der ganzen Herleitung ersichtlich ist, solche Irregularitäten unschädlich.

Will man den bei der Berechnung von Segment oder Exceß vernachlässigten Reihenrest untersuchen, so kann das mit denselben Hilfsmitteln bewerkstelligt werden, die bei der Untersuchung der allgemeinen Interpolationsformel verwendet wurden. Hierbei gelangt man, wie nebenbei bemerkt werden mag, u. a. zu dem Ergebnis, daß die für Segment und Exceß aufgestellten Reihen gewöhnlich nur *semikonvergent* sind.

Sechster Abschnitt.

Numerische Integration: Mittelwertmethoden.

§ 69. Bei der Herleitung der Grundformel für die numerische Integration durch Mittelwerte gehen wir von der Summendefinition des bestimmten Integrals aus. Ist das Flächenstück

$$J = \int_a^b f(x)\, d(x)$$

gegeben, so denke man sich zwischen $x = a$ und $x = b$ auf der Abscissenachse n Punkte eingeschaltet, durch die das ganze Intervall $b-a$ in $n+1$ Teilstrecken zerlegt wird, deren Länge durch $t_0, t_1, \cdots t_n$ bezeichnet werden soll. Innerhalb jeder Teilstrecke suche man aus den zugehörigen Ordinaten je eine aus und bezeichne die ausgewählten Ordinaten mit $y_0, y_1, \cdots y_n$, die zugehörigen Abscissen mit $x_0, x_1, \cdots x_n$. Endlich bilde man das Aggregat

$$K = t_0 y_0 + t_1 y_1 + \cdots + t_n y_n.$$

Dann konvergiert K gegen J, wenn bei unbegrenzt wachsendem n die sämtlichen Teilstrecken gleichzeitig gegen Null konvergieren. Aus diesem Satze folgt, daß der Ausdruck K bei endlichem n zwar nicht den wahren Wert von J, wohl aber eine Annäherung liefert, deren Genauigkeit von der Zahl n, sowie von den ausgewählten Teilstrecken und Ordinaten abhängt.

Schreibt man $K = (b-a)M$, so wird, da die Summe der Teilstrecken t gleich dem Integrationsintervall $b-a$ ist,

$$M = (\sum_q t_q y_q) : (\sum_q t_q), \qquad (q = 0, 1, \cdots n),$$

d. h. die Größe M ist das mit den Gewichten t_q gebildete Mittel der Ordinaten y_q; ferner ist K das Produkt aus diesem Mittel und dem Integrationsintervall. Da es hiernach im wesentlichen nur auf den Mittelwert M ankommt, so wollen wir, um das Intervall $b-a$ nicht immer besonders berücksichtigen zu müssen, die Voraussetzung zu Grunde legen, daß durch geeignete Umformung das gesuchte Integral J

die Grenzen $\pm \frac{1}{2}$ erhalten habe. Das geschieht, wenn man statt x die Veränderliche u durch die Gleichung

$$x = \tfrac{1}{2}(b + a) + (b - a)u$$

einführt und demgemäß

$$\int_a^b f(x)\,dx = (b - a)\int_{-\frac{1}{2}}^{\frac{1}{2}} f\left[\tfrac{1}{2}(b + a) + (b - a)u\right] du$$

schreibt. Das Integrationsintervall für u hat den Wert Eins.

Legt man jetzt das von $-\frac{1}{2}$ bis $+\frac{1}{2}$ genommene Integral

$$J = \int f(x)\,dx$$

zu Grunde, so darf man die gesuchte Näherung K in der Gestalt

$$K = A_0 y_0 + A_1 y_1 + \cdots + A_n y_n$$

schreiben, wo die A gewisse Gewichtsfaktoren bedeuten, durch welche die $n + 1$ dem Integrationsintervall entnommenen Ordinaten y zu einem Mittelwerte verbunden werden; die Summe der A ist hierbei immer gleich Eins.

Die Zahl $n + 1$ gibt an, wie viele Punkte auf der vorgelegten Kurve $y = f(x)$ bei der genäherten Berechnung des Flächeninhaltes J benutzt werden sollen. Ist diese Zahl vorgeschrieben, so ist die Wahl der Kurvenpunkte, oder mit anderen Worten die Festsetzung der zugehörigen Abscissen $x_0, \cdots$ noch offen. Ist diese Wahl irgend wie getroffen, so bleibt für die Gewichte A noch ein gewisser Spielraum übrig, wie aus der geometrischen Entstehung des Ausdruckes K unmittelbar zu entnehmen ist. Diese Umstände führen naturgemäß zu der Frage, ob und wie man die Freiheit, die vorläufig bezüglich der zu verwendenden Kurvenpunkte und Gewichte besteht, dazu ausnutzen könne, um den Fehler der Annäherung, nämlich die Differenz $K - J$ möglichst klein zu machen.

§ 70. Die aufgeworfene Frage setzt voraus, daß man im stande sei, die Werte, die dem Fehler $K - J$ bei verschiedenen für das Mittel K zulässigen Ansätzen zukommen, entweder nach ihrem numerischen Betrage oder doch wenigstens nach ihrer Größenordnung miteinander zu vergleichen. Hieraus geht hervor, daß bezüglich des Verlaufes der Funktion $f(x)$ irgend welche für die gestellte Aufgabe ausreichende Anhaltspunkte gegeben sein müssen, da man andernfalls hinsichtlich des Wertes von $K - J$ völlig im Dunkeln tappen würde. Demgemäß geben wir der betrachteten Frage sogleich eine engere Fassung, indem wir der weiteren Untersuchung die nachstehend aufgeführten Annahmen zu Grunde legen.

Wir setzen zunächst voraus, daß die vorgelegte Funktion $f(x)$ für das ganze Integrationsintervall durch eine konvergente Reihe von der Gestalt

$$f(x) = a_0 g(x)_0 + a_1 g(x)_1 + \cdots \tag{1}$$

darstellbar sei, worin die $g(x)$ völlig bestimmte Funktionen von x, die Größen a dagegen die von der Beschaffenheit der Funktion $f(x)$ abhängenden Koeffizienten bedeuten. Führt man an (1) die vorgeschriebene Integration aus und bezeichnet das über $g(x)_p$ genommene Integral mit G_p, so wird

$$J = a_0 G_0 + a_1 G_1 + \cdots.$$

Setzt man, um Wiederholungen zu vermeiden, fest, daß die Indices $h, k, \cdots$ von 0 bis n, die Indices $p, q, \cdots$ hingegen von 0 bis ∞ laufen, so ist

$$K = \sum_h A_h f(x_h) = \sum_h A_h \sum_p a_p g(x_h)_p,$$

sodaß wir ansetzen dürfen

$$J = \sum_p a_p G_p, \quad K = \sum_p a_p \sum_h A_h g(x_h)_p.$$

Hieraus folgt mit der Abkürzung

$$b_p = \sum_h A_h g(x_h)_p - G_p \tag{2}$$

für den Fehler der Annäherung K der Ausdruck

$$K - J = \sum_p a_p b_p, \tag{3}$$

wo die Faktoren b von den benutzten Werten der Größen x_h und A_h, sowie von der Form der Funktionen $g(x)$ abhängen, hingegen unabhängig sind von der Form der Funktion $f(x)$.

Der vorstehend gemachten Voraussetzung fügen wir nunmehr noch die weitere Annahme hinzu, daß in der Reihe (1) eine rasche Abnahme der Glieder, also eine rasche Konvergenz stattfinde. In diesem Falle sind in der Fehlerreihe (3) die Glieder mit niedriger Nummer die für die Genauigkeit von K gefährlichsten; man wird daher darauf ausgehen, vor allem die niedrigsten Glieder dadurch unschädlich zu machen, daß man die zugehörigen Faktoren b zum Verschwinden bringt. Daraus ergibt sich sofort folgende Lösung der gestellten Aufgabe: man hat, um bei gegebenem n den Fehler $K-J$ möglichst herabzudrücken, nur nötig durch geeignete Wahl der verfügbaren Größen x_h und A_h in der Größenreihe $b_0, b_1, \cdots$ möglichst viele Anfangsglieder zum Verschwinden zu bringen.

Ist die Zahl n, wie wir immer annehmen wollen, vorgeschrieben, so sind vorläufig die $n+1$ Abcissen x_h und die $n+1$ Gewichte A_h,

im ganzen also $2n+2$ Größen verfügbar, sodaß man sich vornehmen kann, ebensoviele „Nullbedingungen“ für die Größen b, nämlich die Gleichungen

$$b_0 = 0, \quad b_1 = 0, \quad \cdots \quad b_{2n+1} = 0 \tag{4}$$

zu erfüllen und damit den Fehlerausdruck auf die Gestalt

$$K - J = a_{2n+2} b_{2n+2} + \cdots$$

zu bringen. Hierbei wird die Lösung, falls die gleichzeitige Erfüllung der Bedingungen (4) überhaupt möglich ist, von der Form der Gleichungen (4), also von der Form der Funktionen $g(x)$ abhängen. Daraus ist zu entnehmen, daß sich im allgemeinen die Wertsysteme der Größen x_h und A_h ändern werden, wenn man die für $f(x)$ zu Grunde gelegte Reihenentwicklung abändert. Infolgedessen ist für die gestellte Aufgabe eine unbegrenzte Mannigfaltigkeit von Lösungen zu erwarten, von denen allerdings nur wenige für den gewöhnlichen Gebrauch in Betracht kommen werden.

§ 71. Um numerisch bestimmte Rechenvorschriften zu gewinnen, haben wir jetzt die zu Grunde gelegte allgemeine Entwicklung

$$f(x) = a_0 g(x)_0 + a_1 g(x)_1 + \cdots$$

in geeigneter Weise zu spezialisieren. Hierbei sollen nur zwei Fälle, nämlich die Potenzreihe und die trigonometrische Reihe, berücksichtigt werden, da andere Formen bisher ohne praktische Bedeutung geblieben sind. Setzt man demgemäß zunächst

$$f(x) = a_0 + a_1 x + a_2 x^2 + \cdots \tag{5}$$

und integriert zwischen den Grenzen $\pm\frac{1}{2}$, so besitzen in der integrierten Reihe

$$J = a_0 G_0 + a_1 G_1 + a_2 G_2 + \cdots$$

die G mit gerader Nummer den Wert

$$G_p = 1 : (p+1)\, 2^p, \tag{6}$$

während die ungeraden G verschwinden. Ferner nehmen die in (4) angesetzten Nullbedingungen unter Berücksichtigung von (2) die Gestalt

$$\sum_h A_h x_h^p = G_p \tag{7}$$

an; es handelt sich also nunmehr um die Frage, wie viele von diesen Bedingungen wirklich erfüllt werden können.

Da in den Gleichungen (7) die Gewichte A nur linear auftreten, so kann man, wie auch die $n+1$ Kurvenpunkte vorgeschrieben sein mögen, durch passende Wahl der Gewichte stets erreichen, daß die

$n+1$ niedrigsten Gleichungen in (7) erfüllt sind, daß also der Fehler von K in der Gestalt

$$K - J = a_{n+1} b_{n+1} + a_{n+2} b_{n+2} + \cdots \tag{8}$$

erscheint. Teilt man z. B. das Integrationsgebiet in drei gleiche Teile und legt die Punkte x_h in die Grenzen der drei Teilstrecken, so ist

$$x_0 = -3:6, \quad x_1 = -1:6, \quad x_2 = 1:6, \quad x_3 = 3:6.$$

Damit findet man die vier ersten Gleichungen des Systems (7) in der Gestalt

$$A_0 + A_1 + A_2 + A_3 = 1, \qquad -3A_0 - A_1 + A_2 + 3A_3 = 0,$$
$$9A_0 + A_1 + A_2 + 9A_3 = 3, \qquad -27A_0 - A_1 + A_2 + 27A_3 = 0.$$

Die beiden Gleichungen rechter Hand liefern zunächst

$$A_3 - A_0 = 0, \qquad A_2 - A_1 = 0$$

oder $A_0 = A_3$, $A_1 = A_2$, womit die übrigen beiden Gleichungen in

$$2A_0 + 2A_1 = 1, \qquad 18A_0 + 2A_1 = 3$$

übergehen. Daraus folgt $A_0 = A_3 = 1:8$, $A_1 = A_2 = 3:8$.

Sind wie im vorliegenden Beispiele die Punkte x_h symmetrisch zum Nullpunkte verteilt, so ist $x_{n-h} = -x_h$. Beachtet man nun, daß in (7) die ungeraden G Null sind, so stellt sich, wenn x_h mit $-x_{n-h}$ und gleichzeitig A_h mit A_{n-h} vertauscht wird, das Gleichungssystem (7) wieder her. Demnach sind bei symmetrischer Verteilung der Punkte x_h die symmetrisch zu den Enden gelegenen Gewichte A_h und A_{n-h} einander gleich.

§ 72. Legt man durch die $n+1$ gegebenen Punkte der Kurve $y = f(x)$ eine Parabel n^{ter} Ordnung mit der Gleichung $y = F(x)$ und unterwirft die Parabel den für die f-Kurve aufgestellten Operationen J und K, so wird bei der soeben eingeführten Gewichtsbestimmung der Fehler $K - J$ für die Parabel in aller Strenge Null, denn in der Fehlerreihe (8) verschwinden jetzt alle Glieder, weil ja bei der Reihe für $F(x)$ die a-Koeffizienten oberhalb der Nummer n sämtlich null sind. Setzt man nun, um die *Lagrange*sche Darstellung benutzen zu können,

$$T(x) = C(x - x_0)(x - x_1)\cdots(x - x_n)$$

und führt die Ableitung $T'(x)$ nebst den Ausdrücken

$$U(x, x_h) = T(x) : [T'(x_h)(x - x_h)]$$

ein, so ist

$$F(x) = \sum_h y_h U(x, x_h),$$

$$J = \sum_h y_h \int U(x, x_h)\, dx, \qquad K = \sum_h y_h A_h.$$

Da nun die y_h unabhängig von den x_h irgendwelche Werte haben können, und da ferner für die Parabel beständig $K = J$ wird, so gelangt man durch Vergleichung der vorstehenden Ausdrücke für K und J zu einer neuen Darstellung der Gewichte, nämlich zu der Formel

$$A_h = \int U(x, x_h)\, dx, \tag{9.a}$$

wofür man auch

$$A_h = \int \prod_k \frac{x - x_k}{x_h - x_k}\, dx \tag{9.b}$$

schreiben kann, wenn man beachtet, daß in dem nach k genommenen Produkt der Index $k = h$ auszufallen hat.

Ob man die Gewichte durch Auflösung der $n + 1$ niedrigsten Gleichungen in (4) oder aus den Integralen (9) ermittelt, ist an sich natürlich gleichgültig. Jedoch mag erwähnt werden, daß die Rechnung nach (9) die bequemere ist, sobald die x nicht unmittelbar bekannt, sondern als die Wurzeln eines gegebenen Polynoms $T(x)$ definiert sind. Bezüglich der hierbei anzuwendenden Kunstgriffe verweise ich auf die Abhandlung von *Gauß* „Methodus nova integralium valores per approximationem inveniendi, (Gesammelte Werke, Band III)“.

Da für die Parabel $y = F(x)$ die Größen K und J übereinstimmen, und da auf der anderen Seite der Wert von K für die Parabel und die Kurve $y = f(x)$ derselbe ist, so beruht die für die f-Kurve aufgestellte Näherung K offenbar darauf, daß man zur Ausführung der Integration an die Stelle der vorgelegten Kurve die Parabel n^{ter} Ordnung setzt, welche durch die $n + 1$ gegebenen Kurvenpunkte $(x_h y_h)$ hindurchgeht.

§ 73. Bei der vorstehenden Betrachtung ist die Verfügung über die Abscissen x_h noch offen geblieben. Wählt man nun die Punkte x_h so, daß sie das Integrationsgebiet in n gleiche Teile zerlegen, so gelangt man zu den Formeln, die zuerst von *Cotes* in seiner „Harmonia mensurarum“ für die numerische Integration aufgestellt worden sind. Es ist dann, wie leicht zu sehen,

$$x_h = (2h - n) : 2n, \qquad (h = 0, 1, 2, \cdots n). \tag{10}$$

Wegen der symmetrischen Anordnung der x_h sind die Gewichte A_h und A_{n-h} einander gleich. Infolgedessen verschwinden auch die Fehlerfaktoren

$$b_p = \sum A_h x_h^p - G_p,$$

sobald p ungerade ist. Demnach fängt die in (8) angesetzte Fehlerreihe

$$K - J = a_{n+1} b_{n+1} + \cdots$$

für $n = 2s$ mit dem Gliede $a_{2s+2} b_{2s+2}$ an, für $n = 2s + 1$ dagegen mit dem Gliede $a_{2s+2} b_{2s+2}$, also beidemal mit demselben a Koeffizienten. Folglich ist der Fortschritt in der Genauigkeit der Annäherung, den man bei ständig wachsendem n zu erwarten hat, im allgemeinen größer, wenn man von einem ungeraden n zu dem nächsten geraden übergeht, als in dem entgegengesetzten Falle. Demgemäß ist es in der Regel vorteilhaft, nur gerade Werte von n zu benutzen.

Die Gewichte finden sich bei *Cotes* und bei *Gauß* (Methodus nova Artikel 4) bis $n = 10$ zusammengestellt. Dieser Zusammenstellung sind die nachstehenden Zahlen für die geraden n entnommen, wobei N den gemeinsamen Nenner der zu einem n gehörigen Gewichte bedeutet.

$$n = 2, \qquad N = 6,$$
$$NA_0 = NA_2 = 1, \qquad NA_1 = 4.$$

$$n = 4, \qquad N = 90,$$
$$NA_0 = NA_4 = 7, \qquad NA_1 = NA_3 = 32, \qquad NA_2 = 12.$$

$$n = 6, \qquad N = 840,$$
$$NA_0 = NA_6 = 41, \quad NA_1 = NA_5 = 216, \quad NA_2 = NA_4 = 27, \quad NA_3 = 272.$$

$$n = 8, \qquad N = 28350,$$
$$NA_0 = NA_8 = 989, \qquad NA_1 = NA_7 = 5888, \qquad NA_2 = NA_6 = -928,$$
$$NA_3 = NA_5 = 10496, \qquad NA_4 = -4540.$$

$$n = 10, \qquad N = 598752,$$
$$NA_0 = NA_{10} = 16076, \quad NA_1 = NA_9 = 106300, \quad NA_2 = NA_8 = -48525,$$
$$NA_3 = NA_7 = 272400, \quad NA_4 = NA_6 = -260550, \quad NA_5 = 427368.$$

Nach den vorstehenden Zahlen treten bei $n = 8$ und $n = 10$ negative Gewichte auf. *Gauß* gibt auch für die einzelnen n den niedrigsten von Null verschiedenen b-Faktor an. Danach ist

n	b_{n+2}
2	$+ 1 : 120$
4	$+ 1 : 2688$
6	$+ 1 : 38880$
8	$+ 37 : 1736\,1504$
10	$+ 26927 : 136\,5000\,0000.$

Diese Zahlen sind klein genug, es ist jedoch zu beachten, daß der Betrag des Fehlers $K - J$ ebensosehr von den a-Koeffizienten wie von den Faktoren b abhängt, daß man also mit den vorstehenden Zahlen nur

dann etwas anfangen kann, wenn man über die ungefähre Größe der a-Koeffizienten unterrichtet ist.

§ 74. Die oben angeführten *Cotes*schen Gewichte liefern bei äquidistanten Ordinaten den vorteilhaftesten Mittelwert, sobald $f(x)$ innerhalb des Integrationsgebietes durch eine rasch konvergierende Potenzreihe dargestellt ist. Trifft diese Voraussetzung nicht zu, so verliert die ganze Beweisführung ihr Fundament, und es kann die unkritische Benutzung der *Cotes*schen Zahlen zu stark fehlerhaften Ergebnissen führen. In dieser Beziehung sind namentlich die Fälle $n = 8$ und $n = 10$ lehrreich, weil bei ihnen negative Gewichte auftreten. Um das deutlich zu machen, diene folgendes Beispiel. Die Funktion $f(x)$ sei zwischen den Grenzen $x = \pm 0.5$ beständig positiv und besitze an den Stellen $x = \pm 0.3$ je ein starkes rasch ansteigendes und ebenso rasch wieder abfallendes Maximum, während die übrigen Funktionswerte auf kleine Beträge beschränkt sein sollen. Dann führt das mit den *Cotes*schen Gewichten für $n = 10$ berechnete Mittel auf einen negativen Wert, während der wahre Wert des Integrals sicher positiv ist. Daraus ist zu entnehmen, daß die Parabel zehnter Ordnung für die vorgelegte Kurve als Annäherung unzureichend ist.

Schwierigkeiten der eben erwähnten Art können sehr leicht eintreten, wenn die zu integrierende Kurve mehrfach zwischen Maximum und Minimum schwankt. Der in solchen Fällen gewöhnlich benutzte Ausweg besteht darin, daß man das Integrationsgebiet in Abschnitte zerlegt und die zugehörigen Teilintegrale einzeln nach der Vorschrift von *Cotes* behandelt. Hierbei werden offenbar die Teilbogen der vorgelegten Kurve durch Bogenstücke ersetzt, die verschiedenen Parabeln angehören, während das ursprüngliche Verfahren für die ganze vorgelegte Kurve eine einzige Parabel höherer Ordnung substituiert.

Zerlegt man z. B. das ganze Integrationsgebiet in m gleiche Abschnitte und behandelt jeden Abschnitt nach der Vorschrift für drei Ordinaten, so kommen im ganzen $2m + 1$ Ordinaten $y_0, y_1, \cdot \cdot y_{2m}$ zur Verwendung, die nach den Gruppen

$$(y_0 y_1 y_2), \quad (y_2 y_3 y_4), \quad (y_4 y_5 y_6), \cdot \cdot$$

geordnet zu den einzelnen Abschnitten die Ordinatenmittel

$$(y_0 + 4y_1 + y_2) : 6, \quad (y_2 + 4y_3 + y_4) : 6, \cdot \cdot$$

liefern. Die vorstehenden Mittel sind, um den zugehörigen Flächeninhalt zu geben, noch mit dem Faktor $1 : m$ zu multiplizieren, weil ja die Breite der einzelnen Abschnitte nicht 1 sondern $1 : m$ ist. Fügt man nunmehr die m Abschnitte zusammen, so wird die ganze Fläche

gleich dem Mittel, das aus den $2m + 1$ Ordinaten mit den Gewichten

$$A_0 = A_{2m} = 1 : 6m,$$
$$A_1 = A_3 = \cdots = A_{2m-1} = 4 : 6m,$$
$$A_2 = A_4 = \cdots = A_{2m-2} = 2 : 6m$$

gebildet wird.

Das vorstehende Gewichtssystem ist unter dem Namen der „*Simpsonschen Regel*“ bekannt und wird häufig vorgezogen, weil die Gewichte bequeme Zahlenwerte besitzen.

§ 75. Nach Darlegung der Methode von *Cotes* kehren wir wieder zu der Voraussetzung zurück, daß über die $n + 1$ Abscissen $x_0, \cdots x_n$ frei verfügt werden dürfe. Da auch für diesen Fall in der Fehlerreihe

$$K - J = a_0 b_0 + a_1 b_1 + \cdots$$

die Faktoren b_p von $p = 0$ bis $p = n$ durch die eingeführte Gewichtsbestimmung jedesmal zum Verschwinden gebracht werden, so entsteht die Frage, ob man nicht durch passende Wahl der Abscissen auch noch $n + 1$ weitere Glieder, nämlich von $p = n + 1$ bis $p = 2n + 1$ beseitigen könne, um dadurch eine erhöhte Genauigkeit von K zu erzielen. Es wird sich zeigen, daß dies in der Tat möglich ist.

Bezeichnet man mit $F[f(x)]$ den Fehler $K - J$, der entsteht, wenn man auf die Funktion $f(x)$ die vorliegende Mittelwertmethode mit $n + 1$ Ordinaten anwendet, so wird, wie wir gesehen haben, $F[f]$ bei beliebiger Wahl der Abscissen in aller Strenge Null, falls man für $f(x)$ eine ganze Funktion setzt, deren Ordnung die Zahl n nicht übersteigt. Kann man nun, wie wir jetzt annehmen wollen, durch passende Wahl der Abscissen auch die Faktoren b_p von $p = n + 1$ bis $p = 2n + 1$ zum Verschwinden bringen, so wird $F[f]$ auch dann noch verschwinden, wenn man für $f(x)$ eine ganze Funktion nimmt, deren Ordnung nicht über $2n + 1$ hinausgeht. Demgemäß wollen wir jetzt unter $f(x)$ ein beliebiges Polynom verstehen, dessen Ordnung höchstens $2n + 1$ ist, und führen weiter das Polynom

$$T(x) = C(x - x_0)(x - x_1) \cdots (x - x_n)$$

ein, dessen Wurzeln die gesuchten Abscissen x_h sind. Dividiert man $f(x)$ durch $T(x)$, so erhält man einen Quotienten $U(x)$ nebst einem Rest $V(x)$, und es wird

$$f(x) = T(x)\,U(x) + V(x). \tag{11}$$

Hierin sind die Polynome $U(x)$ und $V(x)$ höchstens von der Ordnung n. Da ferner die Koeffizienten in $f(x)$ ganz willkürlich gewählt werden dürfen, so sind auch die Koeffizienten in $U(x)$ und $V(x)$ keiner Beschränkung unterworfen.

Führt man an der Gleichung (11) erst die Integration J, sodann die Mittelbildung K aus und subtrahiert die Ergebnisse voneinander, so entsteht die Gleichung

$$F[f] = F[TU] + F[V].$$

Hierin ist $F[V]$ unter allen Umständen Null, weil die Ordnung von V nicht über n hinausgeht. Ferner verschwindet $F[f]$ der Voraussetzung nach. Man erhält also die Bedingung $F[TU] = 0$. Nun ist die zu TU gehörige Näherung K ein Mittel aus den Werten, welche das Produkt TU für die Abscissen x_h annimmt. Dieses Mittel verschwindet aber, weil der Faktor T für die genannten Abscissen verschwindet. Da hiernach in dem zu TU gehörigen Fehler $K - J$ der Bestandteil K Null ist, so kann der Fehler nicht anders verschwinden, als wenn auch J Null wird. Wir gelangen also zu der Bedingung, daß das über TU genommene Integral Null sein muß. Diese Bedingung zerfällt, weil U ein beliebiges Polynom n^{ter} Ordnung bedeutet, in die $n + 1$ Bedingungen

$$\int x^m T(x)\, dx = 0, \qquad (m = 0, 1, \cdot\cdot n), \tag{12}$$

in denen das Integral zwischen den Grenzen ± 0.5 zu nehmen ist. Das System (12) zieht rückwärts wiederum das Verschwinden des über TU genommenen Integrals nach sich.

Die Untersuchung hat sich nunmehr, wie man sieht, auf die Forderung zugespitzt, erstens ein Polynom $T(x)$ zu finden, das den Bedingungen (12) genügt, und zweitens nachzuweisen, daß die Wurzeln von $T(x)$ sämtlich reell, voneinander verschieden und zwischen den Grenzen ± 0.5 enthalten sind, denn nur wenn alle diese Anforderungen gleichzeitig erfüllt werden, kann man von einer wirklichen Lösung der vorliegenden Aufgabe reden.

§ 76. Um $T(x)$ zu ermitteln, bilden wir die wiederholten Integrale von $T(x)$ und schreiben

$$T_1 = \int T\, dx, \qquad T_2 = \int T_1\, dx, \cdot\cdot\,.$$

Läßt man die Integrale jedesmal mit $x = -\frac{1}{2}$ beginnen, so ist T_p durch die p^{te} Potenz des Binoms $x + \frac{1}{2}$ teilbar, wie man sofort erkennt, wenn man bei der Ausführung der Integrationen für $T(x)$ die Entwicklung nach Potenzen von $x + \frac{1}{2}$ zu Grunde legt. Weiter bilden wir das Aggregat

$$W(x, m) = \frac{x^m}{m!} T_1 - \frac{x^{m-1}}{(m-1)!} T_2 + \cdot\cdot \pm \frac{x}{1!} T_m \mp T_{m+1}, \tag{13}$$

worin m eine der Zahlen $0, 1, \cdots n$ bedeutet. Hieraus folgt, wenn man differentiiert,

$$dW(x, m) : dx = x^m T : m!,$$

also, wenn man wieder zwischen den Grenzen $\pm \frac{1}{2}$ integriert und da bei (12) berücksichtigt,

$$W(\tfrac{1}{2}, m) - W(-\tfrac{1}{2}, m) = 0. \qquad (14)$$

Da die Ausdrücke T_p für $x = -\frac{1}{2}$ verschwinden, so gilt das Gleiche von den Größen $W(x, m)$, d. h. es ist $W(-\frac{1}{2}, m) = 0$ und wegen (14) auch $W(\frac{1}{2}, m) = 0$.

Schreibt man die Ausdrücke (13) einzeln hin, so ist

$$W(x, 0) = T_1,$$

$$W(x, 1) = \frac{x}{1!} T_1 - T_2,$$

$$W(x, 2) = \frac{x^2}{2!} T_1 - \frac{x}{1!} T_2 + T_3,$$

u. s. w.

Setzt man hierin $x = \frac{1}{2}$, so verschwinden die linken Seiten und daraus folgt weiter, daß für dieses Argument auch die Größen $T_1, T_2, \cdots$ verschwinden. Demnach hätte man bei der Bildung der Polynome T_p bis $p = n + 1$ als untere Grenze der Integrationen statt $x = -\frac{1}{2}$ auch den Wert $+\frac{1}{2}$ nehmen können, ohne daß dadurch diese Polynome geändert worden wären. Hieraus ergibt sich, daß T_p außer der p^{ten} Potenz von $x + \frac{1}{2}$ auch die p^{te} Potenz von $x - \frac{1}{2}$ als Teiler enthält.

Das Polynom T_{n+1} ist einerseits von der Ordnung $2n + 2$, andererseits enthält es als Teiler die $(n + 1)^{\text{sten}}$ Potenzen von $2x + 1$ und $2x - 1$, die zusammen ebenfalls einen Ausdruck von der Ordnung $2n + 2$ ergeben. Folglich darf man bei passender Wahl des in

$$T(x) = C(x - x_0)(x - x_1) \cdots (x - x_n)$$

mitgenommenen konstanten Faktors C ansetzen

$$T_{n+1} = (2x + 1)^{n+1}(2x - 1)^{n+1} = (4x^2 - 1)^{n+1}, \qquad (15)$$

woraus sofort $T(x)$ folgt, wenn man $(n + 1)$ mal nach x differentiiert.

Um die Wurzeln von $T(x)$ nachzuweisen benutzen wir folgende Bemerkung. Wenn eine ganze Funktion $f(x)$ für die reellen Argumente $x = a$ und $x = b$ verschwindet, so besitzt sie zwischen diesen Argumenten mindestens ein Maximum oder Minimum, also eine Stelle, an der die Ableitung von $T(x)$ verschwindet.

Die Polynome $T_1, T_2, \cdots T_{n+1}$ besitzen sämtlich die Wurzeln $x = -\frac{1}{2}$ und $x = +\frac{1}{2}$. Nach der eben gemachten Bemerkung besitzt also die Ableitung von T_{n+1}, nämlich das Polynom T_n, zwischen den

Argumenten $\pm \frac{1}{2}$ mindestens eine Wurzel und dazu noch die beiden Wurzeln $\pm \frac{1}{2}$. Daraus folgt dann weiter, daß T_{n-1} als Ableitung von T_n zwischen den Argumenten $\pm \frac{1}{2}$ mindestens zwei distinkte Wurzeln und dazu noch die beiden Wurzeln $\pm \frac{1}{2}$ besitzt. In dieser Weise fortfahrend gelangt man zu dem Satze, daß T_1 die beiden Wurzeln $\pm \frac{1}{2}$ und zwischen ihnen mindestens noch n distinkte Wurzeln besitzt. Hieraus folgt dann zum Schlusse, daß dem Polynom $T(x)$ zwischen den beiden Grenzen $\pm \frac{1}{2}$ mindestens $n + 1$ distinkte Wurzeln zukommen. Da nun $T(x)$ überhaupt nur $n + 1$ Wurzeln besitzen kann, so ergibt sich, daß die Wurzeln von $T(x)$ sämtlich reell, von einander verschieden und zwischen den Grenzen $\pm \frac{1}{2}$ enthalten sind; w. z. b. w.

Die vorstehenden Sätze enthalten die Herleitung des Mittelwertverfahrens, das *Gauß* in der „Methodus nova etc.“ entwickelt hat. Er gibt daselbst auch die Zahlen für die Abscissen x_h und die zugehörigen Gewichte bis $n = 6$. Indem ich bezüglich der Zahlenwerte und der von *Gauß* zur Erleichterung der Rechnung herangezogenen Hilfsmittel auf die genannte Abhandlung verweise, möge hier nur der Fall $n = 1$ seiner Einfachheit wegen hervorgehoben werden.

Für $n = 1$ wird

$$T_{n+1} = (4x^2 - 1)^2 = 16x^4 - 8x^2 + 1 .$$

Daraus folgt, wenn man zweimal differentiiert,

$$T(x) = 192x^2 - 16 = 16(12x^2 - 1),$$

also
$$x_0 = -1 : \sqrt{12}, \quad x_1 = 1 : \sqrt{12} .$$

Wegen der symmetrischen Verteilung der x_h wird $A_0 = A_1$, folglich ist, weil die Summe der A immer gleich Eins ist, $A_0 = A_1 = \frac{1}{2}$. Die hieraus folgende Näherung

$$K = \tfrac{1}{2}(y_0 + y_1)$$

ist offenbar der Inhalt des Trapezes, das entsteht, wenn man statt des Kurvenbogens die Sekante setzt, welche durch die beiden benutzten Kurvenpunkte hindurchgeht. Die Fehlerreihe lautet dann nach den von *Gauß* gegebenen Zahlen

$$K - J = -\tfrac{1}{180} a_4 + \cdots ,$$

während bei dem Verfahren von *Cotes* mit drei Ordinaten

$$K - J = \tfrac{1}{120} a_4 + \cdots$$

wird. Der vorstehende Fall kann offenbar zu einer Variante der *Simpson*schen Regel benutzt werden, indem man das ganze Intervall zunächst in m gleiche Teile zerlegt und dann auf jeden Abschnitt die *Gauß*sche Vorschrift für zwei Ordinaten anwendet.

§ 77. Da wir die Entwicklungen, die aus der *Taylor*schen Reihe entspringen, nunmehr verlassen wollen, so mögen noch einige Bemerkungen allgemeiner Natur Platz finden. Hält man die verschiedenen Vorschriften, die oben zur Darlegung gekommen sind, gegeneinander, so kann es nicht zweifelhaft sein, daß die Summenmethode da wo sie überhaupt zulässig ist, im allgemeinen auch den Vorzug verdient. Zur Empfehlung dient ihr neben der Einfachheit der Rechnungsvorschrift vor allem der Umstand, daß man aus dem Verhalten der Differenzen zu ersehen vermag, ob man die Ordinaten dicht genug genommen hat, um die gewollte Genauigkeit des Resultats verbürgen zu können. Denselben Vorzug teilt das Verfahren der Viereckverbesserung, das jedoch der Regel nach hinter der Summenmethode zurückzustehen hat, weil man bei ihm die Ableitungen für die Endordinaten besonders berechnen muß. Bei den oben entwickelten Mittelwertmethoden ist dagegen ein Urteil über die erlangte Genauigkeit nicht ohne weiteres gegeben und im allgemeinen nur dadurch zu erreichen, daß man die Näherung K für verschiedene aufeinander folgende n rechnet und zusieht, wie sich hierbei die Werte von K verhalten. Hierdurch wird aber die Absicht an Ordinaten zu sparen, die namentlich bei der Methode von *Gauß* den Kern der Sache ausmacht, geradezu vereitelt. Damit steht in Einklang, daß die Mittelwertmethoden bei den Aufgaben des wissenschaftlichen Rechnens seither nur in beschränktem Umfange Verwendung gefunden haben; am häufigsten dürfte noch die *Simpson*sche Regel gebraucht werden. In dem Aufsatze „Über die *Cotes*'ischen Integrationsfaktoren (Berliner Jahrbuch für 1863)“ macht *Encke* bezüglich des *Gauß*schen Verfahrens folgende Bemerkung: „*Gauß* hatte wahrscheinlich größere Hoffnungen auf diese Art der Integration gemacht, als er später verwirklicht fand. Er beschäftigte sich zu der Zeit, als ich unter ihm studierte, mit seiner Abhandlung . . . Die Berechnung der Pallasstörungen, die ebenfalls in dieser Zeit mit großem Eifer von ihm verfolgt wurde, mochte mit diesen Untersuchungen in enger Verbindung stehen. Später hat er, soviel mir bekannt, keinen Gebrauch davon gemacht, wovon der Grund nicht schwer nachzuweisen sein dürfte u. s. w.“

§ 78. Wie früher erwähnt, sollte außer der Potenzreihe auch noch die trigonometrische Reihe in die Untersuchung einbezogen werden. Demgemäß verstehen wir jetzt unter $f(x)$ eine periodische Funktion, welche die Periode 2π besitzt, sodaß beständig

$$f(x + 2\pi) = f(x)$$

ist. Zur Veranschaulichung denken wir uns das Argument x nicht mehr als Abscisse auf einer Geraden, sondern als Bogen auf einem mit dem Halbmesser Eins beschriebenen Kreise von einem gegebenen

Nullpunkte aus abgetragen. Jedem Kreispunkte entsprechen dann zwar unendlich viele, um Vielfache von 2π verschiedene Bogenwerte, aber stets nur ein einziger Funktionswert. Weiter setzen wir fest, daß die Integration stets über den ganzen Kreisumfang gehen soll, weil dieser Fall der für die Anwendungen wichtigste ist. Integriert man nun $f(x)$ von a bis $a+2\pi$, so ist das Integral gleich der Summe der Produkte $f(x)dx$, wo für die dx die sämtlichen Bogenelemente des Kreises zu setzen sind. Das Integral ist demnach unabhängig von der Stelle a, mit der die Integration beginnt. Aus diesem Grunde ist es nicht nötig, im folgenden die Integrationsgrenzen jedesmal besonders ersichtlich zu machen.

Wird das über den Kreisumfang genommene Integral mit $2\pi J$ bezeichnet, so ist, weil 2π die Ausdehnung des Integrationsgebietes angibt, die Größe J ein gewisser Mittelwert aus den Beträgen, welche $f(x)$ längs des Kreises annimmt. Um für J eine Näherung K zu finden, denken wir uns auf dem Kreise n Punkte $x_1, \cdot\cdot x_n$ gegeben, dazu die Funktionswerte aufgesucht und diese vermittelst der Gewichte $A_1, \cdot\cdot A_n$ zu dem Ausdruck

$$K = A_1 f(x_1) + A_2 f(x_2) + \cdot\cdot + A_n f(x_n)$$

verbunden. Es handelt sich dann wie früher darum, durch passende Wahl der Größen x_h und A_h den Fehler $K - J$ möglichst herabzudrücken.

Bei der Lösung der Aufgabe gehen wir jetzt von der für periodische Funktionen am nächsten liegenden Darstellung, nämlich von dem Satze aus, daß das vorgelegte $f(x)$ sich für alle reellen x in eine konvergente Reihe von der Gestalt

$$f(x) = c_0 + (2c_1 \cos x + 2s_1 \sin x) + (2c_2 \cos x + 2s_2 \sin x) + \cdot\cdot \tag{16}$$

entwickeln lasse. Diese Reihe ist nun des bequemeren Rechnens wegen zunächst etwas umzuformen, wobei wir festsetzen, daß die Nummern $p, q, \cdot\cdot$ von $-\infty$ bis $+\infty$ laufen sollen, während die Nummern $h, k, \cdot\cdot$ nur von 1 bis n gehen. Außerdem führen wir neben den Koeffizienten $c_0, c_1, \cdot\cdot$ und $s_1, s_2, \cdot\cdot$ die Größen $c(p)$ und $s(p)$ durch die Bedingung ein, daß für jede Zahl g, die der Reihe $0, 1, 2, \cdot\cdot$ angehört,

$$c(g) = c(-g) = c_g, \qquad s(g) = -s(-g) = s_g \tag{17}$$

sein soll. Dann kann man, da $s(0) = -s(-0) = 0$ wird, die obige Reihe in der Gestalt

$$f(x) = \sum_p c(p) \cos px + \sum_p s(p) \sin px \tag{18}$$

schreiben. Ferner ist

$$0 = \sum_p c(p) \sin px - \sum_p s(p) \cos px, \tag{19}$$

da sich in dieser Reihe die Glieder, die zu entgegengesetzten p gehören, vernichten. Verbindet man nun die Gleichungen (18) und (19) vermittelst der Multiplikatoren 1 und i und setzt zur Abkürzung

$$e(p) = c(p) - is(p),$$

so entsteht die Darstellung

$$f(x) = \sum_p e(p) \exp(pxi), \tag{20}$$

die der weiteren Entwicklung zu Grunde gelegt werden soll.

§ 79. Bezeichnet man das über den Kreis genommene Integral von $\exp(pxi)$ mit $2\pi G(p)$, so ist $G(p)$ gleich Null, abgesehen von dem Falle $p = 0$, für welchen $G(0) = 1$ wird. Demnach liefert die Integration der Reihe (20) die Reihe

$$J = \sum_p e(p) G(p),$$

die sich in Wirklichkeit auf das eine Glied

$$J = e(0) G(0) = c(0) = c_0$$

reduziert. Das gesuchte J fällt also mit dem konstanten Gliede der trigonometrischen Entwicklung zusammen.

Für die Näherung K erhält man

$$K = \sum_h A_h f(x_h) = \sum_h A_h \sum_p e(p) \exp(px_h i),$$

woraus

$$K = \sum_p e(p) \sum_h A_h \exp(px_h i)$$

folgt. Damit nimmt der Fehler die Gestalt

$$K - J = \sum_p e(p) b(p) \tag{21}$$

an, wenn zur Abkürzung

$$b(p) = \sum_h A_h \exp(px_h i) - G(p) \tag{22}$$

gesetzt wird. Macht man nun wie früher die Annahme, daß die Koeffizienten c, s rasch abnehmen, so ist es für die Näherung vorteilhaft, möglichst viele Glieder der Reihe

$$b(0), \quad b(\pm 1), \quad b(\pm 2), \cdots$$

zum Verschwinden zu bringen. Da im ganzen $2n$ Größen verfügbar sind, nämlich die n Punkte x_h und die n Gewichte A_h, so kann man sich vornehmen, $2n$ von den Größen $b(p)$ zu beseitigen. Es wird sich zeigen, daß dies in der Tat möglich ist.

Mit der Abkürzung $z = \exp(xi)$, $z_h = \exp(x_h i)$ bilden wir die ganze Funktion

$$T(z) = (z - z_1)(z - z_2) \cdots (z - z_n) \tag{23}$$

und entwickeln diese in der Gestalt

$$T(z) = T_0 + T_1 z + \cdots + T_n z^n. \tag{24}$$

Ferner ersetzen wir in (22) für den Augenblick p durch $q + g$ und schreiben

$$b(q+g) = \sum_h A_h z_h^{q+g} - G(q+g).$$

Multipliziert man diese Gleichung mit T_g und summiert nach g von 0 bis n, so wird

$$\sum_g T_g b(q+g) = \sum_h A_h z_h^q T(z_h) - \sum_g T_g G(q+g).$$

Da hierin nach (23) die Ausdrücke $T(z_h)$ verschwinden, so entsteht zunächst

$$\sum_g T_g b(q+g) = -\sum_g T_g G(q+g). \tag{25}$$

Diese Gleichung wird zur Bestimmung der Koeffizienten T_g führen.

Setzt man bezüglich der Fehlerreihe (21) zunächst fest, daß die Faktoren $b(p)$ verschwinden, wenn p von $-(n-1)$ bis $n-1$ geht, so folgt aus diesen $2n-1$ Nullbedingungen, daß die linke Seite von (25) jedesmal verschwindet, wenn für q eine Zahl der Reihe

$$q = 1-n, \quad 2-n, \quad \cdots -2, \quad -1 \tag{26}$$

gesetzt wird. Ferner verschwinden wegen des Verhaltens der Größen $G(p)$ auf der rechten Seite von (25) alle Glieder, mit Ausnahme des einen, für welches $q = -g$ ist, sodaß die Gleichung (25) die einfache Gestalt

$$0 = -T_{-q} G(0), \quad (q = 1-n, \ 2-n, \cdots -1)$$

annimmt. Demnach wird, wenn q die in (26) vorgeschriebenen Werte durchläuft,

$$0 = T_{n-1} = T_{n-2} = \cdots = T_2 = T_1,$$

sodaß $T(x)$ nunmehr in der Form

$$T(x) = T_0 + T_n z^n$$

erscheint. Hiernach sind die gesuchten Größen z_h die Wurzeln einer binomischen Gleichung von der Gestalt $z^n = P$. Setzt man, um die reellen und rein imaginären Bestandteile hervorzuheben, für P den Ausdruck $Q + Ri$ und führt statt z wieder den Bogen x ein, so müssen die gesuchten x der Gleichung

$$\cos nx + i \sin nx = Q + Ri$$

genügen, welche sofort in die beiden Gleichungen

$$\cos nx = Q, \qquad \sin nx = R$$

zerfällt. Da hiernach $Q^2 + R^2 = 1$ ist, so darf man setzen

$$Q = \cos na, \qquad R = \sin na,$$

wo a einen vorläufig nicht näher bestimmten Bogen bedeutet, und erhält dann mit der Abkürzung $\alpha = 2\pi : n$ für die x_h die n distinkten Werte

$$x_h = a + (h - 1)\alpha, \qquad (h = 1, 2, \cdots n). \tag{27}$$

Die x sind also die Ecken eines regelmäßigen dem Kreise einbeschriebenen n-Ecks, das von der Ecke $x = a$ aus anfängt.

§ 80. Um nunmehr noch die Gewichte A zu finden benutzen wir folgende Bemerkungen. Berechnet man die Werte der irgendwie gegebenen Funktion $f(x)$ für die Ecken des von der Stelle $x = a$ ausgehenden regelmäßigen n-Ecks und verbindet diese Werte zu einem arithmetischen Mittel M, so soll M kurz als das „*n-Eckmittel von der Stelle a aus*“ bezeichnet werden. Wählt man für den Augenblick als Funktion $f(x)$ den mit einem ganzzahligen p gebildeten Ausdruck $\exp(pxi)$, so erhält man für M die Gleichung

$$nM = \exp(pai) + \exp(pai + p\alpha i) + \cdots + \exp[pai + p(n-1)\alpha i].$$

Die rechte Seite ist eine geometrische Reihe, die sich sofort summieren läßt und den Ausdruck

$$nM = \exp(pai)\,[\exp(pn\alpha i) - 1] : [\exp(p\alpha i) - 1]$$

liefert. Ist p nicht durch n teilbar, so verschwindet rechts der Zähler, während der Nenner sicher von Null verschieden ist. In diesem Falle wird also $M = 0$. Ist dagegen p durch n teilbar, so werden in der für M angesetzten Reihe alle Glieder gleich dem Anfangsgliede und man erhält für M den Ausdruck $\exp(pai)$. Faßt man beide Fälle zusammen und spaltet überdies nach den reellen und rein imaginären Bestandteilen, so entsteht folgender Satz: Das von der Stelle a aus gebildete n-Eckmittel liefert für die Ausdrücke

$$\exp(pxi), \qquad \cos px, \qquad \sin px$$

entweder die Werte

$$\exp(pai), \qquad \cos pa, \qquad \sin pa$$

oder aber die Werte Null, je nachdem p durch n teilbar ist oder nicht.

Führt man jetzt in den oben unter (22) für die Fehlerfaktoren aufgestellten Ausdrücken

$$b(p) = \sum_h A_h \exp(p x_h i) - G(p)$$

die in (27) angegebenen Ecken x_h des von der Stelle a ausgehenden n-Ecks ein und setzt alle A_h gleich $1:n$, so verschwinden nach dem vorhin abgeleiteten Satze die Faktoren

$$b(0), \quad b(\pm 1), \quad b(\pm 2), \quad \cdots b[\pm(n-1)].$$

Demnach führen die $2n-1$ bisher benutzten Nullbedingungen zu dem Ergebnis, daß für die gesuchte Näherung K das von einer gewissen Stelle a ausgehende n-Eckmittel der vorgelegten Funktion $f(x)$ zu setzen ist.

In der gefundenen Lösung tritt die vorläufig unbestimmt gebliebene Größe a auf, da die $2n$ verfügbaren Größen x_h und A_h vorläufig nur $2n-1$ Bedingungen unterworfen worden sind. Es ist also noch zu untersuchen, ob sich die angegebene Näherung K durch angemessene Wahl von a noch weiter verbessern läßt.

§ 81. Benutzt man für die trigonometrische Reihe die ursprüngliche Schreibung

$$f(x) = c_0 + 2c_1 \cos x + 2s_1 \sin x + \cdots$$

und führt die oben betrachtete Mittelbildung aus, so wird, weil c_0 mit dem Integral J zusammenfällt,

$$K - J = (2c_n \cos na + 2s_n \sin na) + (2c_{2n} \cos 2na + 2s_{2n} \sin 2na) + \cdots \quad (28)$$

Sieht man für den Augenblick $K-J$ als Funktion des Bogens a an und integriert nach a über den ganzen Kreisumfang, so wird das Integral Null. Daraus folgt, daß $K-J$ teils positive, teils negative Werte annimmt, denn sonst würde, wenn $K-J$ keine Zeichenwechsel erfährt, das Integral notwendig einen von Null verschiedenen positiven oder negativen Wert besitzen. Da wir es ferner immer nur mit stetigen Funktionen zu tun haben, so muß $K-J$ für gewisse Werte von a durch Null hindurchgehen, d. h. bei passender Wahl von a liefert das Mittel K den strengen Wert J. Leider bleibt diese Bemerkung für gewöhnlich unfruchtbar, weil der nötige Anhalt zur Ermittelung des günstigsten Wertes von a fehlt. Deshalb mag hier nur der besondere Fall berührt werden, wo $f(x)$ eine gerade Funktion von x ist.

Wenn $f(x)$ gerade ist, so fehlen in der trigonometrischen Reihe alle Sinusglieder und man erhält

$$K - J = 2c_n \cos na + 2c_{2n} \cos 2na + \cdots.$$

Soll nun z. B. der Fehler mit dem Gliede der Ordnung $2m$ beginnen, so genügt es $n = m$ zu nehmen, wenn man $a = \pi : 2m$ wählt, während für $a = 0$ die Eckenzahl n gleich $2m$ gemacht werden müßte. Es besteht also zwischen den beiden Werten von a ein erheblicher Unterschied hinsichtlich der Anzahl der nötigen Funktionswerte. Bei näherem Zusehen zeigt sich jedoch, daß der Unterschied in dem erforderlichen Arbeitsaufwande nur unbedeutend ist. Rechnet man nämlich mit dem $2m$-Eck von $a = 0$ aus, so sind, weil $f(x)$ gerade ist, die Funktionswerte des dritten und vierten Quadranten nur eine Wiederholung der Werte aus dem ersten und zweiten Quadranten, man hat also nur $m + 1$ Funktionswerte wirklich zu berechnen, während bei dem m-Eck von $a = \pi : 2m$ aus m Funktionswerte zu berechnen sind. Es wird also bei letzterem Ansatz nur ein Funktionswert gespart. Daraus erklärt sich die Gewohnheit der Rechner, den Ausgangspunkt $a = 0$ zu bevorzugen.

Hält man die verschiedenen oben behandelten Mittelwertmethoden gegeneinander, so fällt sofort auf, wie die Änderung der für $f(x)$ benutzten Reihenentwicklung auch die Rechnungsvorschrift ändert: der Gegenstand ist offenbar durch die beiden hier untersuchten Formen bei weitem nicht erschöpft.

Siebenter Abschnitt.

Trigonometrische Reihen.

§ 82. Die im letzten Abschnitt behandelte Integration einer periodischen Funktion $f(x)$ findet in ausgedehnter Weise Anwendung, wenn es sich darum handelt, die Koeffizienten der Entwicklung

$$f(x) = c_0 + 2c_1 \cos x + 2s_1 \sin x + \cdots$$

numerisch zu bestimmen. Hält man die in § 78 flg. eingeführte Bezeichnung fest und multipliziert die in § 78 (20) angesetzte Gleichung

$$f(x) = \sum_p e(p) \exp(pxi) \tag{1}$$

mit $\exp(-qxi)$, so verschwinden, wenn man nunmehr über den Kreisumfang integriert, rechterhand alle Terme mit Ausnahme des zu $p = q$ gehörigen Gliedes, und man erhält

$$\int f(x) \exp(-qxi)\, dx = 2\pi e(q). \tag{2}$$

Daraus fließen wegen der Beziehung $e(q) = c(q) - is(q)$, wenn man reelles und rein imaginäres trennt, die beiden Gleichungen

$$2\pi c(q) = \int f(x) \cos qx\, dx. \qquad 2\pi s(q) = \int f(x) \sin qx\, dx. \tag{3}$$

Da die zu integrierenden Ausdrücke periodische Funktionen sind, so könnte man zunächst auf den Gedanken kommen, für sämtliche Koeffizienten genäherte Werte abzuleiten, indem man die Methode des n-Eckmittels auf die obigen Integrale anwendet. Eine kurze Überlegung lehrt jedoch, daß man auf diesem Wege höchstens bis zu einem gewissen Punkte gelangen kann. Wenn man nämlich mit einem bestimmten n-Eck rechnet, so ist numerisch weiter nichts gegeben, als das System der n Funktionswerte $f(x_h)$. Jeder dieser Werte liefert eine und nur eine Gleichung von der Form

$$f(x_h) = c_0 + 2c_1 \cos x_h + 2s_1 \sin x_h + \cdots. \tag{3}$$

Man wird also über den Verlauf der Funktion und dem entsprechend über die Werte der Koeffizienten c, s überhaupt nicht mehr aussagen

können, als sich aus den vorstehenden n Gleichungen herausholen läßt, es sei denn, daß man anderswoher noch weitere Aussagen über die Koeffizienten heranzuziehen vermag. Dies wird noch deutlicher werden, wenn man die Mittelbildung wirklich ausführt.

§ 83. Der Einfachheit halber soll das n-Eck fortan von der Stelle $x = 0$ aus konstruiert werden. Es ist das keine wesentliche Einschränkung der Allgemeinheit, wie man sofort erkennt, wenn man statt $f(x)$ die Funktion $g(x) = f(a + x)$ zu Grunde legt. Des weitern multiplizieren wir die in (1) für $f(x)$ angesetzte Darstellung mit $\exp(-qxi)$ und nehmen von der so entstehenden Gleichung

$$f(x)\cos qx - i f(x)\sin qx = \sum_p c(p)\exp(pxi - qxi)$$

das Mittel über das n-Eck, wobei die Mittelwerte von $f(x)\cos qx$ und $f(x)\sin qx$ kurz mit (c, q) und (s, q) bezeichnet werden sollen. Auf der rechten Seite liefern nur diejenigen Glieder einen Beitrag, in denen der Quotient $(p - q) : n$ gleich einer ganzen Zahl r ist. Da ferner wegen $a = 0$ das Mittel aus $\exp(nrxi)$ gleich Eins wird, so erhalten wir unter Berücksichtigung der Bedeutung von $c(p)$ die Beziehung

$$(c, q) - i(s, q) = \sum_r c(q + nr) = \sum_r c(q + nr) - i\sum_r s(q + nr),$$

welche in die beiden Gleichungen

$$(c, q) = \sum_r c(q + nr), \qquad (s, q) = \sum_r s(q + nr) \tag{4}$$

zerfällt. Das vorstehende System liefert zunächst unendlich viele Gleichungen, da ja q irgend eine ganze Zahl bedeutet. Die nähere Betrachtung wird jedoch zeigen, daß man es in Wahrheit nur mit den Wiederholungen eines Systems von n Gleichungen zu tun hat. Bezeichnet man nämlich für den Augenblick die auf den rechten Seiten von (4) auftretenden Reihen mit $C(q)$ und $S(q)$ und denkt sich dieselben ausgeschrieben, so erkennt man ohne Mühe, daß

$$C(q + n) = C(q), \qquad S(q + n) = S(q) \tag{5}$$

ist. Demnach wiederholen sich die Gleichungen (4) periodisch, wenn man q um ein Vielfaches von n wachsen läßt; es reicht also aus, q auf die Zahlen $0, 1, \cdots n - 1$ zu beschränken. Denkt man sich ferner die Reihen für $C(-q)$ und $S(-q)$ ausgeschrieben und beachtet die Relationen

$$c(-p) = c(p), \qquad s(-p) = -s(p),$$

so ergibt sich

$$C(-q) = C(q), \qquad S(-q) = -S(q),$$

woraus nach (5)

$$C(n-q) = C(q), \quad S(n-q) = -S(q) \tag{6}$$

folgt. Hiernach erhält man bereits alle voneinander verschiedenen Gleichungen des Systems (4), wenn man mit q in der Reihe $0, 1, 2, \cdots$ nicht über $\frac{1}{2}n$ hinausgeht.

Ist $n = 2m + 1$, so kommen, da $(s, 0)$ und $S(0)$ identisch verschwinden, aus dem System (4) nur die Gleichungen in Betracht, welche zu den Größen

$$\begin{matrix} (c, 0), & (c, 1), \cdots (c, m), \\ & (s, 1), \cdots (s, m) \end{matrix} \tag{7}$$

gehören. Ist dagegen $n = 2m$, so wird außer $S(0)$ auch $S(m)$ identisch Null, wie sich ergibt, wenn man die Reihe für $S(m)$ ausschreibt. Demnach kommen aus dem System (4) nur die Gleichungen in Betracht, die zu den Größen

$$\begin{matrix} (c, 0), & (c, 1), \cdots (c, m-1), & (c, m), \\ & (s, 1), \cdots (s, m-1) & \end{matrix} \tag{8}$$

gehören. Die Berechnung der Mittel (c, p) und (s, p) liefert also jedesmal gerade n Gleichungen zwischen den Koeffizienten der vorgelegten Reihe.

Setzt man die in Betracht kommenden Gleichungen einzeln an und führt statt der Größen $c(p)$ und $s(p)$ wieder die Koeffizienten c_p und s_p ein, so erhält man aus den Kosinusgrößen

$$\left.\begin{aligned} (c, 0) &= c_0 + (c_n + c_n) + (c_{2n} + c_{2n}) + \cdots, \\ (c, 1) &= c_1 + (c_{n+1} + c_{n-1}) + (c_{2n+1} + c_{2n-1}) + \cdots, \\ (c, 2) &= c_2 + (c_{n+2} + c_{n-2}) + (c_{2n+2} + c_{2n-2}) + \cdots, \\ &\text{u. s. w.} \end{aligned}\right\} \tag{9}$$

wo das Gesetz des Fortschreitens klar ist. Die letzte Gleichung dieses Systems lautet, wenn $n = 2m$ ist,

$$(c, m) = c_m + (c_{n+m} + c_{n-m}) + (c_{2n+m} + c_{2n-m}) + \cdots,$$

wofür man schreiben kann

$$(c, m) = 2c_m + 2c_{3m} + 2c_{5m} + \cdots. \tag{10}$$

Aus den Sinusgrößen erhält man die Gleichungen

$$\left.\begin{aligned} (s, 1) &= s_1 + (s_{n+1} - s_{n-1}) + (s_{2n+1} - s_{2n-1}) + \cdots, \\ (s, 2) &= s_2 + (s_{n+2} - s_{n-2}) + (s_{2n+2} - s_{2n-2}) + \cdots, \\ &\text{u. s. w.} \end{aligned}\right\} \tag{11}$$

wo das Gesetz des Fortschreitens ebenfalls ersichtlich ist.

Die in (9) und (11) angesetzten Systeme von zusammen n Gleichungen enthalten in einer übersichtlichen Form das, was sich über die Koeffizienten c, s auf Grund der gegebenen Funktionswerte und der dazu gehörigen oben in (3) aufgestellten Gleichungen aussagen läßt. Zu dieser Aussage wird nun bei den Anwendungen der Zusatz gemacht, daß die Werte der c, s nur bis zu einer gewissen Stellennummer l hin merklich sind. Wählt man daraufhin n größer als $2l$, so dürfen in (9) und (11) die rechten Seiten auf ihre Anfangsglieder reduziert werden und man erhält

$$\begin{aligned} c_0 &= (c, 0), \quad c_1 = (c, 1), \quad c_2 = (c, 2), \cdots \\ & \qquad\qquad\; s_1 = (s, 1), \quad s_2 = (s, 2), \cdots \end{aligned} \tag{12}$$

wobei nur zu beachten ist, daß für $n = 2m$ nach (10) der Koeffizient c_m nicht gleich (c, m) sondern gleich $\frac{1}{2}(c, m)$ zu setzen ist.

§ 84. Die vorstehenden Formeln führen zu dem Verfahren, das weitaus am häufigsten benutzt wird, wenn es sich darum handelt, für periodische Gebilde oder Vorgänge auf rein numerischem Wege die zugehörige trigonometrische Entwicklung herzustellen. In der Tat zeichnet sich das Verfahren durch Einfachheit und Übersichtlichkeit aus und beansprucht nicht mehr Funktionswerte, als Koeffizienten bestimmt werden. Sieht man für den Augenblick von den nachher zu besprechenden Vereinfachungen ab, so gestaltet sich der Rechnungsgang folgendermaßen.

Nachdem die Teilung, d. h. der Wert von n, festgesetzt ist, bildet man mit $w = 2\pi : n$ die Argumentreihe

$$0, w, 2w, \cdots (n-1)w \tag{13}$$

und legt zugleich die Tabelle an, welche die Vielfachen dieser Argumentreihe soweit als nötig enthält. Daran schließt sich die Berechnung der Funktionswerte, die wir der Reihe (13) entsprechend kurz durch die Symbole (0), (1), $\cdots (n-1)$ bezeichnen wollen, während die zugehörigen Ecken der Teilung einfach durch die Nummern $0, 1, \cdots n-1$ unterschieden werden sollen. Weiter ist das Schema für die Ausführung der Multiplikationen mit seinen Zeilen und Spalten anzulegen, wobei man zweckmäßig die Zeilen den Eckennummern zuordnet. Die der Ecke q entsprechende Zeile enthält dann den Funktionswert $f(qw)$ und die Faktoren

$$\cos qw, \quad \sin qw, \quad \cos 2qw, \quad \sin 2qw, \quad \text{u. s. w.}$$

Für $\cos(0qw)$ ist keine besondere Spalte vorzusehen, da $(c, 0)$ einfach das Mittel aus den Funktionswerten selber ist. Rechnet man logarithmisch, so werden natürlich für die Faktoren sogleich ihre Logarithmen angesetzt. Die Spalten für die Produkte $f(qw)\cos(pqw)$

und $f(qw)\sin(pqw)$ kann man neben den zugehörigen Faktorenspalten aussparen und zwar sogleich doppelt, um für die positiven und negativen Produkte je eine eigene Spalte zur Verfügung zu haben. Sind die Produkte gebildet und summiert, so liefert die Division der Summen durch n bezw. $\frac{1}{2}n$ die Koeffizienten c_0, $2c_1$, $2s_1$, $\cdots$.

Sind die miteinander zu verbindenden Faktoren höchstens dreiziffrig, so kann man überall statt der Logarithmen die Zahlen ansetzen, weil man in diesem Falle mit dem Rechenschieber oder den vorhandenen Produkttafeln ebenso bequem arbeitet, wie mit der Logarithmentafel. Ebenso wird man die Zahlen ansetzen, wenn man die Produkte jeder Zeile mit der Maschine rechnen will, was durchaus nicht unzweckmäßig ist. Nur hat man dann, um die verkettete Multiplikation anwenden zu können, den Faktoren jeder Zeile noch ihre Differenzen hinzuzufügen.

§ 85. Der Rechnungsgang verlangt in der vorstehend beschriebenen Anordnung für jeden Koeffizienten die Berechnung von n Produkten. Diese Zahl läßt sich jedoch reduzieren, weil in jeder Faktorenspalte gewisse Zahlen — vom Vorzeichen abgesehen — wiederholt vorkommen. Infolgedessen kann man eine Reihe von Multiplikationen sparen, indem man statt $AC \pm BC$ sogleich $(A \pm B)C$ rechnet.

Bezeichnet man die von 0 bis $n-1$ laufende Eckennummer allgemein mit g, so erhält man mit den Funktionswerten (0), (1), $\cdots$ für die Größen (c, q) und (s, q) die Gleichungen

$$n(c, q) = \sum_g (g)\cos(qgw), \qquad n(s, q) = \sum_g (g)\sin(qgw). \tag{14}$$

Hierin lassen sich zunächst die Glieder zusammenziehen, deren Ecken symmetrisch zum Nullpunkte liegen, denn zu je zwei symmetrisch gelegenen Ecken gehören gleiche Kosinus und entgegengesetzte Sinus. Demgemäß setzen wir für die Ecken, die dem ersten Halbkreise — unter Ausschluß der Stellen $x = 0^0$ und $x = 180^0$ — angehören, zunächst

$$\sigma(g) = (g) + (n - g) \tag{15.a}$$

und fügen dazu die Bestimmung $\sigma(0) = (0)$ und — für den Fall eines geraden n — die Bestimmung $\sigma(\frac{1}{2}n) = (\frac{1}{2}n)$. Dann wird

$$n(c, q) = \sum_g \sigma(g)\cos(qgw), \tag{15.b}$$

wo die Summation nach g jetzt nur die Ecken des ersten Halbkreises umfaßt. In ähnlicher Weise bilden wir für die (s, q)

$$\delta(g) = (g) - (n - g), \tag{15.c}$$

$$n(s, q) = \sum_g \delta(g)\sin(qgw), \tag{15.d}$$

wo bei der Summation, die wiederum nur über den ersten Halbkreis

geht, die Glieder mit $\delta(0)$ und $\delta(\frac{1}{2}n)$ nicht in Betracht kommen, weil ihre Sinusfaktoren verschwinden.

Die Reihen (15) lassen sich, wenn n ungerade ist, im allgemeinen nicht weiter zusammenziehen: für spezielle Werte von n und q ist das allerdings möglich, bietet jedoch keine erheblichen Vorteile, falls nicht etwa n sehr groß ist. Ist dagegen n gerade, so stimmen für die Ecken, die sich zu 180^0 ergänzen, die Sinus und die Kosinus, vom Vorzeichen abgesehen, überein und lassen eine weitere Zusammenziehung zu. Zunächst werde bei geradem n für die Ecken, die dem ersten Quadranten — unter Ausschluß der Stelle $x = 90^0$ — angehören

$$\sigma\sigma(g) = \sigma(g) + \sigma(\tfrac{1}{2}n - g), \qquad \delta\sigma(g) = \sigma(g) - \sigma(\tfrac{1}{2}n - g) \qquad (16.a)$$

gesetzt, mit dem Hinzufügen, daß, wenn $n = 4m$ ist,

$$\sigma\sigma(\tfrac{1}{4}n) = \sigma(\tfrac{1}{4}n)$$

sein soll. Damit läßt sich, wenn man die geraden und ungeraden q getrennt behandelt, die Reihe (15.b) in der Gestalt

$$n(c, q) = \sum_g \sigma\sigma(g)\cos(qgw), \qquad (q \text{ gerade}), \qquad (16.b)$$

$$n(c, q) = \sum_g \delta\sigma(g)\cos(qgw), \qquad (q \text{ ungerade}) \qquad (16.c)$$

schreiben, wo die Summation nur über den ersten Quadranten geht, und die Größe $\delta\sigma(g)$ für die Ecke $gw = 90^0$ herausfällt, weil der zugehörige Kosinus verschwindet. Entsprechend bildet man für die Größen (s, q)

$$\sigma\delta(g) = \delta(g) + \delta(\tfrac{1}{2}n - g), \qquad \delta\delta(g) = \delta(g) - \delta(\tfrac{1}{2}n - g) \qquad (16.d)$$

und fügt dazu, wenn $n = 4m$ ist, die Bestimmung

$$\sigma\delta(\tfrac{1}{4}n) = \delta(\tfrac{1}{4}n).$$

Daraus ergibt sich

$$n(s, q) = \sum_g \delta\delta(g)\sin(qgw), \qquad (q \text{ gerade}), \qquad (16.e)$$

$$n(s, q) = \sum_g \sigma\delta(g)\sin(qgw), \qquad (q \text{ ungerade}), \qquad (16.f)$$

wo bei der Summation, die wieder nur über den ersten Quadranten geht, die Ecke $g = 0$ und — für $n = 4m$ und für ein gerades q — auch die Ecke $gw = 90^0$ ausfällt, weil die zugehörigen Sinusfaktoren verschwinden.

§ 86. Die Größen $\sigma(g), \cdots$, deren Ermittelung sich unmittelbar an die Berechnung der Funktionswerte (g) anzuschließen hat, können nach einer einfachen Regel gebildet werden. Ist n ungerade, so gruppiert man die (g) in zwei Zeilen; die erste Zeile enthält die

Ecken des ersten Halbkreises, darunter stehen in der zweiten Zeile die Ergänzungen zum Vollkreise. So entsteht z. B. für $n = 9$ das Zeilenpaar

$$\begin{matrix} (0) & (1) & (2) & (3) & (4) \\ - & (8) & (7) & (6) & (5). \end{matrix}$$

Die beiden Zeilen sind durch Addition und Subtraktion zu verbinden, wobei der Strich unter (0) für die Addition als eine Null gilt, für die Subtraktion dagegen das Ausfallen der Ecke (0) anzeigt. Auf diese Weise entstehen aus den beiden Zeilen die Größen

$$\begin{matrix} \sigma(0) & \sigma(1) & \sigma(2) & \sigma(3) & \sigma(4) \\ & \delta(1) & \delta(2) & \delta(3) & \delta(4). \end{matrix}$$

Ist n von der Form $2m + 2$, so gruppiert man in vier Zeilen. Die erste Zeile enthält die Ecken des ersten Quadranten, darunter setzt man in der dritten Zeile erst die Ergänzungen zu 180° und darauf unter die hingeschriebenen Ecken in der zweiten und vierten Zeile die Ergänzungen zum Vollkreise. Auf diese Weise entsteht z. B. für $n = 10$ die Anordnung

$$\begin{matrix} (0) & (1) & (2) \\ - & (9) & (8) \\ (5) & (4) & (3) \\ - & (6) & (7) \end{matrix}$$

wo die Striche unter (0) und (5) dieselbe Rolle zu spielen haben, wie im vorigen Falle. Werden die Zeilenpaare erst durch Addition und dann durch Subtraktion verbunden, so entstehen die neuen Zeilen

$$\begin{matrix} \sigma(0) & \sigma(1) & \sigma(2) \\ \sigma(5) & \sigma(4) & \sigma(3) \\ \hline & \delta(1) & \delta(2) \\ & \delta(4) & \delta(3) \end{matrix}$$

aus denen durch nochmalige Addition und Subtraktion die Größen

$$\begin{matrix} \sigma\sigma(0) & \sigma\sigma(1) & \sigma\sigma(2) \\ \delta\sigma(0) & \delta\sigma(1) & \delta\sigma(2) \\ & \sigma\delta(1) & \sigma\delta(2) \\ & \delta\delta(1) & \delta\delta(2) \end{matrix}$$

entspringen. Wenn $n = 4m$ ist, so gruppiert man nach derselben Regel. Beispielsweise erhält man für $n = 12$ erst die Zeilen

$$\begin{matrix} (0) & (1) & (2) & (3) \\ - & (11) & (10) & (9) \\ (6) & (5) & (4) & \\ - & (7) & (8) & \end{matrix}$$

darauf

$$\begin{array}{cccc} \sigma(0) & \sigma(1) & \sigma(2) & \sigma(3) \\ \sigma(6) & \sigma(5) & \sigma(4) & - \\ & \delta(1) & \delta(2) & \delta(3) \\ & \delta(5) & \delta(4) & - \end{array}$$

und darauf

$$\begin{array}{cccc} \sigma\sigma(0) & \sigma\sigma(1) & \sigma\sigma(2) & \sigma\sigma(3) \\ \delta\sigma(0) & \delta\sigma(1) & \delta\sigma(2) & \\ & \sigma\delta(1) & \sigma\delta(2) & \sigma\delta(3) \\ & \delta\delta(1) & \delta\delta(2) & \end{array}$$

Wenn n durch 4 teilbar ist, so lassen sich die Produktsummen noch weiter nach Oktanten zusammenziehen, jedoch hat man dann die Werte von q noch weiter zu spezialisieren, sodaß die Anzahl der zu berücksichtigenden Formeln vermehrt und dadurch die Übersichtlichkeit der Rechnung beeinträchtigt wird.

Sind die Größen $\sigma(q)$, $\cdot\cdot$ ermittelt, so verläuft die weitere Rechnung in derselben Weise, wie vorhin in § 84 für die unreduzierten Reihen angegeben worden ist, nur daß jetzt das Multiplikationsschema in zwei oder vier Abschnitte mit verminderter Zeilen- und Spaltenanzahl zerfällt. Die Menge der zu berechnenden Produkte sinkt hierbei für ein ungerades n auf fast die Hälfte, für ein gerades n auf fast ein Viertel des ursprünglichen Betrages. Aus diesem Grunde werden, wenn nicht besondere Umstände im Wege sind, allgemein die Teilungen mit gradem n vorgezogen. Da ferner für ein durch 4 teilbares n die Ecke $x = 90^0$ mit den bequemen Multiplikatoren $0, +1, -1$ auftritt, so wird die Teilung von der Form $n = 4m$ mit besonderer Vorliebe benutzt.

Da der Rechnungsgang bis zur Ermittelung der Koeffizienten c, s keine besonderen Kontrollen bietet, so ist es wünschenswert für die fertige Entwicklung eine Prüfung zu haben. Solche Proben ergeben sich, wenn man mit der fertigen Reihe rückwärts wieder einige der gegebenen Funktionswerte berechnet. So erhält man für $x = 0$ die Gleichung

$$(0) = f(0) = c_0 + 2c_1 + 2c_2 + \cdots,$$

durch welche die c-Koeffizienten kontrolliert werden. Für die s-Koeffizienten läßt sich eine Gleichung von ähnlicher Einfachheit nicht aufstellen, dagegen kann man die Gleichung

$$\delta(1) = f(w) - f(-w) = 4s_1 \sin w + 4s_2 \sin 2w + \cdots$$

benutzen.

§ 87. Die hier entwickelte Rechnung mit dem regelmäßigen n-Eck setzt voraus, daß die Glieder, die bei dem Übergange von den

Gleichungen (9) und (11) zu dem System (12) unterdrückt worden sind, auch tatsächlich unmerklich seien, da andernfalls die errechneten Werte der Koeffizienten um den Betrag der unterdrückten Glieder fehlerhaft werden. Ebenso wird, wenn die mit den gefundenen Koeffizienten gebildete endliche Reihe als eine genügende numerische Darstellung der vorgelegten Funktion $f(x)$ dienen soll, vorausgesetzt, daß die Gesamtheit der fortgelassenen Glieder oder kurz der „Rest" der unendlichen Reihe $f(x)$ nichts merkliches ausmache. Nun ist man häufig nicht in der Lage schon vor Beginn der Rechnung überschlagen zu können, wie groß man die Eckenzahl n zu wählen habe, um der Erfüllung jener Voraussetzungen sicher zu sein. In solchen Fällen ist man daher darauf angewiesen, die Zahl n vorläufig nach Gutdünken zu wählen, mit dem Vorbehalte, die Rechnung nachträglich zu vervollständigen, wenn sich das gewählte n schließlich als zu klein herausstellt. Bei dieser Vervollständigung stehen zwei Wege offen, die der Reihe nach besprochen werden sollen.

Ist das ursprüngliche n nicht sehr groß, so empfiehlt es sich in der Regel als das Einfachste, eine neue Bestimmung der Koeffizienten auszuführen, bei der sogleich die doppelte Eckenzahl zu Grunde gelegt wird. Hierbei lassen sich, wie unmittelbar ersichtlich ist, die in der alten Rechnung vorkommenden Funktionswerte und Produkte ohne weiteres in die neue Rechnung herübernehmen. Wäre also z. B. die Rechnung mit einem 12-Eck begonnen worden, so hätte die verlangte Vervollständigung durch den Übergang auf ein 24-Eck zu erfolgen. Wollte man statt dessen ein 16-Eck benutzen, so würde man nichts gewinnen, selbst wenn die Teilung mit $n = 16$ an sich ausreichend wäre, denn man hätte, wie leicht zu sehen, die Werte von $f(x)$ und die zugehörigen Produkte für 12 neue Ecken zu berechnen und könnte überdies von der alten Rechnung nur einen kleinen Teil wieder heranziehen; ebenso ungünstig oder noch ungünstiger stünde die Sache bei jeder anderen zwischen 12 und 24 gelegenen Eckenzahl.

§ 88. Die Eckenverdoppelung wird unzweckmäßig, wenn n groß ist, und wenn man überdies aus dem Verhalten der bereits gefundenen Koeffizienten erkennt, daß nur noch wenige Glieder der Entwicklung nachzutragen sind. In diesem Falle ist ein anderer Weg vorteilhafter, den man als eine *Zusammensetzung von Vielecken* bezeichnen kann. Das hierbei innezuhaltende Verfahren läßt sich am raschesten an einem Beispiel deutlich machen. Angenommen die ursprünglich mit einem 12-Eck durchgeführte Rechnung solle durch Hinzufügung von vier weiteren Koeffizienten vervollständigt werden. Dann genügt es, wie sich zeigen wird, für die zweite Rechnung ein Vieleck zu wählen, das mindestens vier Ecken besitzt, die von den Ecken des 12-Ecks verschieden sind, wie das z. B. bei dem 8-Eck der Fall ist. Bezeichnet

man nämlich wie oben mit (c, p) und (s, p) die Produktmittel für das 12-Eck, so gelten, wenn man die Glieder oberhalb der Stellennummer 8 fortläßt, nach (9) die Gleichungen

$$\left.\begin{array}{l}(c, 0)=c_0, \quad (c, 1)=c_1, \quad (c, 2)=c_2, \quad (c, 3)=c_3, \\ (c, 4)=c_4+c_8, \quad (c, 5)=c_5+c_7, \quad (c, 6)=2c_6,\end{array}\right\} \tag{17}$$

und nach (11) die Gleichungen

$$\left.\begin{array}{lll}(s, 1)=s_1, & (s, 2)=s_2, & (s, 3)=s_3, \\ (s, 4)=s_4-s_8, & (s, 5)=s_5-s_7. &\end{array}\right\} \tag{18}$$

Auf dieselbe Weise entstehen, wenn die zum 8-Eck gehörigen Mittel mit $[c, p]$ und $[s, p]$ bezeichnet werden, die Gleichungen

$$\left.\begin{array}{lll}[c, 0]=c_0+2c_8, & [c, 1]=c_1+c_7, & [c, 2]=c_2+c_6, \\ [c, 3]=c_3+c_5, & [c, 4]=2c_4, &\end{array}\right\} \tag{19}$$

und weiter

$$[s, 1]=s_1-s_7, \quad [s, 2]=s_2-s_6, \quad [s, 3]=s_3-s_5. \tag{20}$$

Von den zwölf in (17) und (19) enthaltenen Gleichungen für die c-Koeffizienten sind drei eine Folge der übrigen; es ist nämlich

$$\begin{gathered}(c, 0)+2(c, 4)=[c, 0]+[c, 4], \\ (c, 1)+(c, 3)+(c, 5)=[c, 1]+[c, 3], \\ 2(c, 2)+(c, 6)=2[c, 2].\end{gathered}$$

Diese Beziehungen, die auch noch bestehen bleiben, wenn man für die Größen (c, p) und $[c, p]$ die vollständigen Reihen nach (4) einsetzt, lassen erkennen, daß man in (19) die Gleichungen mit $[c, 2]$, $[c, 3]$ und $[c, 4]$ streichen kann, daß also nur die Berechnung von $[c, 0]$ und $[c, 1]$ wirklich nötig ist. Man erhält dann die gesuchten c-Koeffizienten, indem man die Formeln

$$\begin{gathered}c_0=(c, 0), \quad c_1=(c, 1), \quad c_2=(c, 2), \quad c_3=(c, 3), \\ 2c_4=2(c, 4)+(c, 0)-[c, 0], \quad c_5=(c, 5)-(c, 1)-[c, 1], \\ 2c_6=(c, 6), \quad c_7=[c, 1]-(c, 1), \quad 2c_8=[c, 0]-(c, 0)\end{gathered}$$

durchrechnet.

Behandelt man die Gleichungen (18) und (20) in ähnlicher Weise, so kann man zunächst wegen der Relation

$$(s, 1)-(s, 3)+(s, 5)=[s, 1]-[s, 3]$$

die Gleichung für $[s, 3]$ streichen und erhält aus den sieben übrigbleibenden Gleichungen unter Vernachlässigung von s_8 die Ausdrücke

$$s_1 = (s, 1), \quad s_2 = (s, 2), \quad s_3 = (s, 3),$$
$$s_4 = (s, 4), \quad s_5 = (s, 1) - (s, 5) = [s, 1],$$
$$s_6 = (s, 2) = [s, 2], \quad s_7 = (s, 1) = [s, 1].$$

Das vorstehende Beispiel läßt zur Genüge erkennen, wie sich bei anderen Vielecken die Zusammensetzung gestalten würde. Die Gleichungssysteme, die hierbei auftreten, sind von so einfacher Beschaffenheit, daß ihre Auflösung keinerlei Schwierigkeiten bereitet. Im übrigen ist jedoch das Verfahren immer nur als eine Aushilfe anzusehen: so liefern in dem obigen Beispiel die benutzten sechszehn Ecken ein weniger günstiges Ergebnis, als man mit dem regelmäßigen 16-Eck erhalten würde, denn in den Ausdrücken für die Sinuskoeffizienten mußte man bei der Zusammensetzung s_8 vernachlässigen, während bei dem regelmäßigen 16-Eck erst s_9 unterdrückt wird.

Die vor Beginn der Rechnung häufig bestehende Ungewißheit über die Anzahl der mitzunehmenden Glieder kann recht lästig werden, wenn man eine Anzahl von Funktionen gleichzeitig zu bearbeiten hat, deren Entwicklungen ungleich rasch konvergieren. Man wird dadurch zu der Frage geführt, ob es nicht möglich sei ein Verfahren auszuarbeiten, bei dem alles Umhertasten fortfällt und die Rechnung schrittweise jedesmal nur soweit getrieben wird, als einerseits notwendig, andererseits hinreichend ist. Solche Vorschriften lassen sich in der Tat aufstellen; da sie jedoch ohne besondere Hilfstafeln nur unbequem zu verwenden sind, so beschränke ich mich darauf, den Grundgedanken zu skizzieren und schicke zu dem Ende eine Bemerkung über die Auflösung linearer Systeme voraus.

§ 89. In dem Gleichungssystem

$$\left.\begin{aligned} Y_1 &= (1, 1)X_1 + (1, 2)X_2 + (1, 3)X_3 + \cdots, \\ Y_2 &= (2, 1)X_1 + (2, 2)X_2 + (2, 3)X_3 + \cdots, \\ &\text{u. s. w.} \end{aligned}\right\} \qquad (21)$$

sei die Anzahl der Größen X, Y und entsprechend die Anzahl der Gleichungen nicht weiter begrenzt, während die Koeffizienten (p, q) durch das dafür aufgestellte Bildungsgesetz numerisch vollständig bestimmt und derart beschaffen sein sollen, daß man die n ersten Gleichungen des Systems jedesmal nach den n vordersten X auflösen kann; es wird verlangt, die Auflösung so anzulegen, daß die ganze für ein bestimmtes n erforderliche Rechnung benutzbar bleibt, wenn man nachträglich auf einen größeren Wert von n übergehen will. Zu dem Ende bilde man zunächst mit gewissen Multiplikatoren $(p)_q$ die Verbindungen

$$\left.\begin{aligned} Z_1 &= Y_1(1)_1, \\ Z_2 &= Y_1(2)_1 + Y_2(2)_2, \\ Z_3 &= Y_1(3)_1 + Y_2(3)_2 + Y_3(3)_3, \\ &\text{u. s. w.} \end{aligned}\right\} \quad (22)$$

und bestimme die Multiplikatoren so, daß in Z_p die Größen X_1 bis X_{p-1} fehlen und X_p den Koeffizienten Eins erhält. Dadurch entstehen die Bedingungen

$$\begin{aligned} 1 &= (1, 1)(1)_1, \\ 0 &= (1, 1)(2)_1 + (2, 1)(2)_2, \\ 1 &= (1, 2)(2)_1 + (2, 2)(2)_2, \\ 0 &= (1, 1)(3)_1 + (2, 1)(3)_2 + (3, 1)(3)_3, \\ 0 &= (1, 2)(3)_1 + (2, 2)(3)_2 + (3, 2)(3)_3, \\ 1 &= (1, 3)(3)_1 + (2, 3)(3)_2 + (3, 3)(3)_3, \\ &\text{u. s. w.} \end{aligned}$$

Sind die Multiplikatoren $(p)_q$ aus den vorstehenden Bedingungen ermittelt, so erscheinen die Z in der Gestalt

$$\left.\begin{aligned} Z_1 &= X_1 + [1, 2]X_2 + [1, 3]X_3 + [1, 4]X_4 + \cdots, \\ Z_2 &= \qquad\quad\; X_2 + [2, 3]X_3 + [2, 4]X_4 + \cdots, \\ Z_3 &= \qquad\qquad\qquad\quad\; X_3 + [3, 4]X_4 + \cdots, \\ &\text{u. s. w.} \end{aligned}\right\} \quad (23)$$

wo der Koeffizient von X_q in Z_p durch den Ausdruck

$$[p, q] = (1, q)(p)_1 + (2, q)(p)_2 + (3, q)(p)_3 + \cdots$$

bestimmt ist, der von selber bei dem Gliede mit $(p)_p$ abbricht. Aus dem System (23) ergeben sich nun die X durch schrittweises Eliminieren in der Gestalt

$$\left.\begin{aligned} X_1 &= Z_1 + Z_2[1]_2 + Z_3[1]_3 + Z_4[1]_4 + \cdots, \\ X_2 &= \qquad Z_2 \quad + Z_3[2]_3 + Z_4[2]_4 + \cdots, \\ X_3 &= \qquad\qquad\quad Z_3 \quad + Z_4[3]_4 + \cdots, \\ &\text{u. s. w.} \end{aligned}\right\} \quad (24)$$

wo die zu X_p gehörigen Multiplikatoren $[p]_q$ der Reihe nach aus den Gleichungen

$$\begin{aligned} 0 &= [p, p+1] + [p]_{p+1}, \\ 0 &= [p, p+2] + [p+1, p+2]\cdot[p]_{p+1} + [p]_{p+2}, \\ 0 &= [p, p+3] + [p+1, p+3]\cdot[p]_{p+1} + [p+2, p+3]\cdot[p]_{p+2} + [p]_{p+3}, \\ &\text{u. s. w.} \end{aligned}$$

zu bestimmen sind.

Soll die Auflösung für die n ersten Gleichungen unter Fortlassung der X oberhalb der Nummer n erfolgen, so kommt dies darauf hinaus, daß man die nicht zu berücksichtigenden X einfach gleich Null

setzt. In diesem Falle sind die Z bis zur Nummer n zu bilden, während die folgenden Z von selber Null werden. Infolgedessen brechen die in (24) gegebenen Ausdrücke von selbst mit Z_n ab. Will man nun nachträglich die Rechnung auf $n+r$ Gleichungen mit ebensovielen X erweitern, so bleibt die alte Rechnung unverändert bestehen und erfährt nur folgende Zusätze: 1) Hat man der Reihe der Z noch die Größen für die Nummern von $n+1$ bis $n+r$ hinzuzufügen; 2) sind in den Ausdrücken der bereits berechneten X noch die Glieder für die neuen Z nachzutragen; 3) sind mit dem neuen Z die Ausdrücke für die neuen X zu berechnen.

Die beschriebene Auflösung ist offenbar hinlänglich bequem, sobald die Koeffizienten (p, q) des vorgelegten Systems ein für allemal gegeben und daraufhin die Werte der Multiplikatoren $(p)_q$ und $[p]_q$ tabuliert worden sind.

§ 90. Nach der vorstehenden Einschaltung nehmen wir die Betrachtung der trigonometrischen Reihe wieder auf. Wir denken uns jetzt auf dem Kreise außer dem Nullpunkte eine unbegrenzte Reihe von Punkten gegeben, die paarweise symmetrisch zum Nullpunkte liegen und in einer bestimmten Reihenfolge geordnet sind. Außerdem soll das Bildungsgesetz der Reihe, die in der Gestalt

$$0, \quad \pm x_1, \quad \pm x_2, \quad \pm x_3, \cdots \tag{25}$$

geschrieben werden möge, so eingerichtet sein, daß keine Wiederholungen auftreten, daß also kein späteres Punktepaar mit einem früheren zusammenfällt. Solche Reihen kann man auf unendlich viele Arten ansetzen; bezeichnet z. B. w einen Bogen, der zum Vollkreise inkommensurabel ist, so würde die Reihe

$$0, \quad \pm w, \quad \pm 2w, \quad \pm 3w, \cdots \tag{26}$$

der gestellten Bedingung genügen.

Bildet man nun aus der Funktion

$$f(x) = c_0 + 2c_1 \cos x + 2s_1 \sin x + \cdots$$

die beiden neuen Funktionen

$$g(x) = f(x) + f(-x) = 2c_0 + 4c_1 \cos x + 4c_2 \cos 2x + \cdots,$$
$$u(x) = f(x) - f(-x) = 4s_1 \sin x + 4s_2 \sin 2x + \cdots,$$

so ist $g(x)$ eine gerade, $u(x)$ eine ungerade Funktion von x. Denkt man sich ferner $g(x)$ für die Punkte der Reihe (25), d. h. für die Ecken $x_0 = 0$, x_1, x_2, $\cdots$ berechnet, so entsteht ein Gleichungssystem von der Form

$$g(x_h) = 2c_0 + 4c_1 \cos x_h + \cdots, \quad (h = 0, 1, 2, \cdots). \tag{27}$$

Dieses läßt sich nach dem vorhin angegebenen Verfahren behandeln, wobei den Größen X die gesuchten c Koeffizienten, den Größen Y die Funktionswerte $g(x_h)$, endlich den Koeffizienten (p, q) die Faktoren der c-Größen in (27) entsprechen. Bei der Ausführung berechnet man dann die $g(x_h)$ vorläufig bis zu einer nach Gutdünken gewählten Nummer $h = n$ hin und bildet daraus zuerst die Größen Z nach (22). Besitzen in dieser Z-Reihe die letzten Glieder merkliche Werte, so ist daraus zu entnehmen, daß die c-Koeffizienten gleicher Nummer ebenfalls noch merklich sind. Man wird dann in der Reihe der $g(x_h)$ um einige Glieder über die ursprünglich angenommene Nummer n hinaus weiter gehen, dazu die entsprechenden Z-Größen ermitteln und hiermit solange fortfahren, bis die letzten Z unmerklich geworden sind. Hierauf berechnet man nach (24) die X, d. h. die c-Koeffizienten und hat damit die erste Hälfte der Arbeit erledigt. Die zweite Hälfte besteht darin, daß man das Gleichungssystem

$$u(x_h) = 4s_1 \sin x_h + 4s_2 \sin 2x_h + \cdots \qquad (28)$$

in genau derselben Weise behandelt.

Die wesentlichen Stücke des beschriebenen Rechnungsganges sind offenbar darin zu sehen, daß man auf die gleichmäßige Verteilung der benutzten Ecken verzichtet und auf Grund einer im voraus ein für allemal festgesetzten unbegrenzten Eckenreihe die Hilfsgrößen für die schrittweise Auflösung der Gleichungen tabuliert; ohne eine solche Tabulierung würde die Rechnung natürlich erheblich zeitraubender ausfallen, als bei der Benutzung eines regelmäßigen Vielecks.

Die Rechnung mit einem unregelmäßigen Vieleck ist zuerst von *Leverrier* vorgeschlagen und auch benutzt worden. Die dazu gehörigen Vorschriften findet man in der „Addition III" zu „Annales de l'Observatoire Impérial de Paris Vol. I". *Leverrier* benutzt, was für das Prinzip der Methode unwesentlich ist, eine von den obigen Festsetzungen etwas abweichende Eckenreihe, die dadurch entsteht, daß man einen zum Vollkreise inkommensurabelen Bogen w vom Nullpunkte ausgehend beständig in derselben Richtung hin abträgt. Dieselbe Eckenreihe wird auch in dem Aufsatze von *Encke* „Über die Entwicklung einer Funktion in eine periodische Reihe nach Herrn *Leverriers* Vorschlag (Berliner Jahrbuch für 1860)" zu Grunde gelegt. Später wurde der Gegenstand wieder aufgenommen von *Hoüel* in „Mémoire sur le développement des fonctions en séries périodiques au moyen de l'interpolation (Annales de l'Observatoire Impérial de Paris, Vol. VIII)". *Hoüel* benutzt, was eine Verbesserung war, die in (26) mit einem inkommensurabelen Bogen w angesetzte Eckenreihe und leitet für die Gleichungen vier verschiedene Auflösungsmethoden ab, darunter auch die oben behandelte. In der Folgezeit ist dann das Verfahren, soviel ich sehen kann, ganz in den Hintergrund ge-

treten, wobei anscheinend auch rein äußerliche Gründe mitgewirkt haben. Gleichwohl dürfte es der Mühe wert sein, dasselbe noch weiter auszubilden, wobei man besondere Rücksicht auf die Sicherheit der Rechnung zu nehmen hätte. Was es damit für eine Bewandtnis hat, wird aus den nachstehenden Bemerkungen erhellen.

§ 91. Reduziert man die Reihe für $f(x)$ auf die m niedrigsten Glieder und bestimmt diese Glieder aus den Funktionswerten $y_h = f(x_h)$, die zu m irgendwie verteilten Ecken $x_1, \cdots x_m$ gehören, so erscheinen die Koeffizienten als lineare Verbindungen der y in der Gestalt

$$y_1 K_1 + y_2 K_2 + \cdots + y_m K_m, \tag{29}$$

wo die Faktoren K durch die gegebenen Ecken vollständig bestimmt sind. Ebenso erscheint auch $f(x)$, wenn man die für die Koeffizienten gefundenen Ausdrücke einsetzt, als eine lineare Verbindung der y in der Gestalt

$$f(x) = y_1 U_1 + y_2 U_2 + y_m U_m, \tag{30}$$

wo die U endliche trigonometrische Reihen bedeuten, deren Koeffizienten durch die Faktoren K gegeben sind. Da nun die y im allgemeinen mit gewissen Abrundungsfehlern behaftet sind, so ist es für die Sicherheit der Rechnung wesentlich, daß die Faktoren K und folgeweise auch die Ausdrücke U auf mäßige Werte beschränkt bleiben, weil andernfalls jene Abrundungsfehler den Wert von $f(x)$ erheblich fälschen können.

Hat man ein *regelmäßiges* m-Eck zu Grunde gelegt, so läßt sich die Frage nach der Sicherheit der Rechnung sogleich befriedigend beantworten, denn die Faktoren K sind dann, wie sich aus den früheren Entwicklungen ergibt, von dem Typus $\cos x : n$ oder $\sin x : n$, also niemals größer als $1 : n$. Infolgedessen gehen die Fehler der einzelnen y verkleinert in den Ausdruck für $f(x)$ ein, und es hat deshalb auch einen guten Sinn, bei der Rechnung in den Koeffizienten von $f(x)$ eine oder auch zwei Decimalen mehr mitzunehmen, als in den gegebenen y. Ist hingegen das benutzte m-Eck *unregelmäßig*, und liegen überdies zwei Ecken sehr nahe beieinander, so kommen, wie hier nicht weiter bewiesen werden soll, unter den Faktoren K große Zahlen vor, die zunächst die Berechnung der Koeffizienten von $f(x)$ unsicher machen. Ferner kommen auch unter den Werten der Ausdrücke U große Beträge vor, denn da die Koeffizienten der U sich in der Form

$$\frac{1}{2\pi}\int U \cos px\, dx \quad \text{oder} \quad \frac{1}{2\pi}\int U \sin px\, dx$$

darstellen lassen, so können große Koeffizienten in einem U nicht anders auftreten, als wenn dieses U längs des Kreises mindestens an

einer Stelle groß wird. Das enge Zusammenrücken zweier Ecken macht also auch die aus der ermittelten Reihe berechneten Werte von $f(x)$ unsicher.

Die letzte Bemerkung führt unmittelbar zu der Frage, um die es sich handelt. Will man bei der schrittweisen Koeffizientenbestimmung nach dem oben entwickelten Verfahren eine symmetrische unbegrenzte Eckenreihe von der Form

$$0, \quad \pm x_1, \quad \pm x_2, \quad \pm x_3, \ldots$$

zu Grunde legen, so hat man das Bildungsgesetz dieser Reihe derart einzurichten, daß jederzeit, also für drei, für fünf, für sieben u. s. w. Ecken, der Einfluß der den gegebenen y anhaftenden Abrundungsfehler in $f(x)$ möglichst heruntergedrückt wird. Diese Forderung führt, indem man zuerst mit einem Dreieck rechnet, zur Festlegung von x_1. Ist x_1 festgelegt, so führt dieselbe Forderung, für das Fünfeck erhoben, zu der Festlegung von x_2, und so geht es weiter. Ich unterlasse es diese vielleicht nicht ganz leichte Aufgabe hier weiter zu untersuchen, da es mir nur darauf ankam nachzuweisen, daß der Gegenstand der weiteren Verfolgung wert ist, denn der auf *Leverrier* zurückzuführende Gedanke, mit der Eckenzahl schrittweise nur bis zu dem jedesmal notwendigen und hinreichenden Werte zu gehen, ist als durchaus rationell zu bezeichnen.

§ 92. Will man nach Abschluß der Rechnung prüfen, ob die merklichen Glieder der trigonometrischen Entwicklung tatsächlich mit der gewünschten Vollständigkeit berücksichtigt worden seien, so kann dazu folgende Überlegung dienen.

Bedeutet $F(x)$ die Summe der mit den errechneten Koeffizienten gebildeten endlichen Reihe, so beruhen die gefundenen Werte der Koeffizienten auf der Bedingung, daß in den benutzten Ecken $F(x)$ mit $f(x)$ übereinstimme. Diese Übereinstimmung wird auch noch außerhalb jener Ecken vorhanden sein, wenn die bei der Herleitung von $F(x)$ unberücksichtigt gebliebenen Glieder der unendlichen Reihe $f(x)$ in ihrer Gesamtheit unmerklich sind. Dagegen muß, wenn der unterdrückte Reihenrest nicht unmerklich ist, außerhalb der Ecken die Differenz $F(x) - f(x)$ merkliche Werte annehmen. Berechnet man also für einen oder mehrere Punkte, die etwa in der Mitte zwischen benachbarten Ecken liegen, einerseits direkt den Wert von $f(x)$, andererseits die Summe der Reihe $F(x)$, so gewähren die Beträge der Differenz $F - f$ einen Anhalt, wie weit die Reihe $F(x)$ für den Verlauf der vorgelegten Funktion $f(x)$ genügt.

§ 93. Die bisher behandelte Rechnung mit dem regelmäßigen oder unregelmäßigen Vieleck geht davon aus, daß die Anzahl der zu

Grunde gelegten Funktionswerte nicht größer als die Anzahl der ermittelten Koeffizienten ist. Hierbei ist stillschweigend die Auffassung leitend, daß die Berechnung der Funktionswerte einen erheblichen Teil des ganzen Arbeitsaufwandes darstelle und daß es deshalb angezeigt sei, mit den Funktionswerten zu sparen. Es kann nun aber auch vorkommen, daß der Verlauf von $f(x)$ bequem tabuliert oder graphisch gegeben ist, daß also die Ansetzung des einzelnen Funktionswertes nur eine ganz untergeordnete Mühe verursacht. Man kann dann fragen, ob sich nicht aus diesem Umstande Vorteil ziehen lasse, indem man mehr Gleichungen ansetzt, als Unbekannte gesucht werden. Das ist in der Tat möglich. Bisher ist allerdings nur einziges hierher gehöriges Verfahren ausgearbeitet worden, das jetzt behandelt werden soll. In der Arbeit

> *P. Tschebychef*, Sur l'interpolation dans le cas d'un grand nombre de données fournies par les observations. (Mémoires de l'Académie des sciences de St. Pétersbourg Série VII, Tome I),

die auch, mit Beispielen und Hilfstafeln versehen, in der Abhandlung

> *P. Harzer*, Über eine von Herrn *Tschebychef* angegebene Interpolationsformel. (Astronomische Nachrichten Band 115)

reproduziert ist, wird für eine beliebig vorgelegte Funktion eine algebraische Reihenentwicklung gesucht, die gewissen, auf den ersten Anblick ziemlich verwickelt aussehenden Nebenbedingungen genügen soll. Geht man jedoch der Sache näher nach, so ergibt sich, daß man in Wahrheit eine eigenartige Methode der trigonometrischen Entwicklung vor sich hat, und daß die algebraische Einkleidung nur eine fremde Hülle darstellt, die den Kern des Verfahrens verdeckt und die ganze Herleitung unnütz erschwert.

Als Ausgangspunkt soll hier die nachstehend beschriebene Operation dienen. Indem wir das Argument der periodischen Funktion $f(x)$ wieder als Bogen auf dem Einheitskreise abtragen, konstruieren wir, von der Stelle $x = a$ aus vorwärts gehend, ein regelmäßiges $2n$-Eck und ordnen den einzelnen Ecken, mit der Ecke $x = a$ beginnend, die Nummern $0, 1, \cdot\cdot 2n - 1$ zu. Weiter berechnen wir für die geraden Ecken die Werte von $f(x)$, für die ungeraden Ecken hingegen die Werte von $-f(x)$ und bilden aus diesen $2n$ Zahlen das arithmetische Mittel. Das Ergebnis der ganzen Operation soll für den Augenblick kurz mit (a, n) bezeichnet werden.

Wählt man als Funktion $f(x)$ zunächst den mit einem ganzzahligen p gebildeten Ausdruck $\exp(pxi)$ und setzt zur Abkürzung

$$\pi : 2n = h, \qquad \exp(2phi) = z,$$

so wird im vorliegenden Falle

$$2n(a, n) = \exp(ai)\,[1 - z + z^2 - \cdot\cdot - z^{2n-1}], \tag{31}$$

woraus

$$2n(a, n) = \exp(ai)(1 - z^{2n}) : (1 + z)$$

folgt. Der Zähler des Bruches auf der rechten Seite ist beständig Null, folglich kann (a, n) nur dann einen von Null verschiedenen Wert annehmen, wenn der Nenner ebenfalls verschwindet, wenn also $z = -1$ wird. In diesem Falle muß der Quotient $p : n$ gleich einer ungeraden Zahl k oder $p = kn$ sein, womit dann aus der Reihe (31) unmittelbar $(a, n) = \exp(pai)$ folgt. Trennt man nunmehr nach den reellen und den rein imaginären Bestandteilen, so entsteht folgender Satz: Die Operation (a, n), an den Ausdrücken $\cos px$ und $\sin px$ ausgeführt liefert, die Werte

$$(a, n) = \cos pa \quad \text{und} \quad (a, n) = \sin pa,$$

sobald der Quotient $p : n$ eine ungerade ganze Zahl k ist; für die übrigen Werte der ganzen Zahl p ist (a, n) beständig null.

§ 94. Nach dieser Vorbereitung führen wir jetzt die beschriebene Operation an der trigonometrischen Entwicklung von $f(x)$ aus, wobei mit einer kleinen Abänderung der früheren Bezeichnungen die Schreibung

$$\begin{aligned} f(x) = A(0) + A(1)\cos x + A(2)\cos 2x + \cdots \\ + B(1)\sin x + B(2)\sin 2x + \cdots \end{aligned} \tag{32}$$

benutzt werden soll. Es liefern dann nur diejenigen Glieder einen Beitrag, deren Nummer ein ungerades Vielfaches von n ist, und man erhält

$$(a, n) = \sum_k [A(kn)\cos(kna) + B(kn)\sin(kna)], \quad (k = 1, 3, 5, \cdots).$$

Für die numerischen Anwendungen sollen nur die beiden Werte $a = 0$ und $a = h$ benutzt und die entsprechenden Werte von (a, n) mit $U(n)$ und $V(n)$ bezeichnet werden. Es wird dann

$$U(n) = A(n) + A(3n) + A(5n) + \cdots, \tag{33.a}$$

$$V(n) = B(n) - B(3n) + B(5n) - \cdots. \tag{33.b}$$

Mit Hilfe dieser Gleichungen sind jetzt die A, B durch die als bekannt vorausgesetzten Größen U, V auszudrücken. Für die Berechnung der U, V gilt dabei folgende einfache Regel. Man teilt den Kreisumfang vom Nullpunkte aus in $4n$ gleiche Teile und ermittelt die $4n$ Funktionswerte von $f(x)$ für die Teilpunkte. Weiter summiert man die gefundenen Funktionswerte getrennt für die vier aus den Teilpunkten zu bildenden regelmäßigen n-Ecke, die mit den Stellen

$$x = 0, \quad x = \pi : 2n, \quad x = 2\pi : 2n, \quad x = 3\pi : 2n$$

anfangen. Bezeichnet man dann diese vier Summen mit I, II, III, IV, so ist

$$U(n) = (\mathrm{I} - \mathrm{III}) : 2n, \quad V(n) = (\mathrm{II} - \mathrm{IV}) : 2n. \tag{34}$$

Ist die Funktion $f(x)$ gerade oder ungerade, so fällt, wie aus (33) zu entnehmen ist, die Berechnung der Summen II, IV oder I, III fort, weil dann nur die A oder B gebraucht werden.

Da bei den vorstehenden Betrachtungen die Zahl n mindestens gleich Eins ist, so kommt in den Gleichungen (33) der Koeffizient $A(0)$ nicht vor; er wird jedoch in einfacher Weise gefunden, wenn man für die größten Werte von n, die bei der ganzen Rechnung auftreten, die Ausdrücke

$$A(0) = (\mathrm{I} + \mathrm{II} + \mathrm{III} + \mathrm{IV}) : 4n$$

bildet.

§ 95. Um die Gleichungen (33) aufzulösen ist zunächst ein Hilfssatz einzuschalten. Die Nummer k durchlaufe die Reihe der ungeraden natürlichen Zahlen $1, 3, 5, \cdots$; in dieser Reihe streiche man die durch 3 teilbaren Zahlen und bezeichne die stehen bleibenden Nummern mit k_3; ferner streiche man von den k_3 die durch 5 teilbaren Zahlen und bezeichne die stehen bleibenden Nummern mit k_5; in derselben Weise bilde man die Nummernreihen $k_7, k_{11}, \cdots$, indem man der Reihe nach die aufeinander folgenden Primzahlen heranzieht. Des weitern sei $F(x)$ eine Funktion von der Art, daß die Reihe

$$F(1) + F(2) + F(3) + \cdots$$

unbedingt konvergiert, und es werde der Ausdruck

$$u(h) = F(h) + F(3h) + F(5h) + \cdots = \sum_k F(kh) \tag{35}$$

für alle ungeraden Zahlen $h = 1, 3, 5, \cdots$ angesetzt; dann läßt sich das so entstehende Gleichungssytem durch ein einfaches Verfahren nach den Größen F auflösen. Zu dem Ende bilde man aus den $u(h)$ neue Ausdrücke $u(h)_3, u(h)_5, \cdots$, deren Indices durch die aufeinander folgenden Primzahlen gegeben sind, nach folgendem unmittelbar verständlichen Schema

$$\begin{aligned}
u(h)_3 &= u(h) - u(3h) = \sum F(kh) - \sum F(3kh) = \sum F(k_3 h),\\
u(h)_5 &= u(h)_3 - u(5h)_3 = \sum F(k_3 h) - \sum F(5k_3 h) = \sum F(k_5 h),\\
u(h)_7 &= u(h)_5 - u(7h)_5 = \sum F(k_5 h) - \sum F(7k_5 h) = \sum F(k_7 h),
\end{aligned}$$
u. s. w.

wo die Summen sich immer auf das hinter dem Zeichen $\sum$ stehende k beziehen. Die rechter Hand stehenden Reihen reduzieren sich, wenn der Prozeß unbegrenzt fortgetzt wird, schließlich auf das Anfangs-

glied $F(h)$. Ferner ergibt sich das Bildungsgesetz der u-Größen ohne Mühe, wenn man den Anfang der geforderten Rechnung wirklich ausführt. Es wird nämlich der Reihe nach

$$\begin{aligned}
u(h)_3 &= u(h) - u(3h),\\
u(5h)_3 &= u(5h) - u(15h),\\
u(h)_5 &= u(h) - u(3h) - u(5h) + u(15h),\\
u(7h)_5 &= u(7h) - u(21h) - u(35h) + u(105h),\\
u(h)_7 &= u(h) - u(3h) - u(5h) - u(7h)\\
&\quad + u(15h) + u(21h) + u(35h) - u(105h),
\end{aligned}$$

u. s. w.

Wie man sieht treten auf der rechten Seite jedesmal nur solche $u(kh)$ auf, in denen die Zahl k ein Produkt von lauter ungleichen Primfaktoren ist; ferner ist das Vorzeichen positiv oder negativ, je nach dem die Anzahl der ungleichen Primfaktoren gerade oder ungerade ist. Demgemäß können wir nunmehr übersichtlicher schreiben

$$F(h) = \sum_k a(k)\, u(kh), \tag{36}$$

wenn unter $a(k)$ die Zahl 0 oder $+1$ oder -1 verstanden wird, jenachdem k durch das Quadrat einer Primzahl teilbar ist oder aber nur ungleiche Primfaktoren in gerader oder ungerader Menge enthält.

Auf dieselbe Weise läßt sich auch die Lösung des Gleichungssystems

$$u(h) = F(h) - F(3h) + F(5h) - \cdots \tag{37}$$

finden, das sich von (35) durch die alternierenden Vorzeichen unterscheidet. Aus der für ungerade h, k geltenden Gleichung

$$i^{(h-1)(k-1)} = 1$$

folgt, wenn man den Exponenten ausmultipliziert,

$$i^{hk+1} = i^{h+k}. \tag{38}$$

Damit bilde man aus (37) der Reihe nach

$$i u(h) = \sum_k i^k F(kh),$$

$$i^{h+1} u(h) = \sum_k i^{h+k} F(kh) = \sum_k i^{hk+1} F(kh),$$

und setze

$$i^{h+1} u(h) = v(h), \qquad i^{h+1} F(h) = G(h).$$

Dann entsteht das Gleichungssystem

$$v(h) = \sum_k G(kh),$$

das dieselbe Gestalt wie (35) besitzt und zu der Lösung

$$G(h) = \sum_k a(k)\, v(kh)$$

führt, aus der weiter

$$i^{h+1} F(h) = \sum_k a(k)\, i^{hk+1} u(kh)$$

und wegen (38)

$$F(h) = \sum_k a(k)\, i^{k-1} u(kh)$$

folgt. Demnach ist die Lösung von (37) durch die Gleichungen

$$F(h) = \sum_k b(k)\, u(kh), \quad b(k) = a(k)\, i^{k-1} \tag{39}$$

gegeben.

§ 96. Mit den vorstehenden Formeln läßt sich jetzt die Auflösung der Gleichungen (33) ohne weiteres hinschreiben. Versteht man unter n eine feste gerade oder ungerade Zahl, während die ungeraden Zahlen h, k beweglich bleiben, so entstehen aus (35) durch die Annahme

$$u(h) = U(hn), \quad F(h) = A(hn)$$

Gleichungen, die dem System (33. a) angehören. Demgemäß ergibt sich nach (36), wenn nunmehr $h = 1$ gesetzt wird,

$$A(n) = \sum_k a(k)\, U(kn). \tag{40}$$

Ebenso folgen aus (37) mit der Annahme

$$u(h) = V(hn), \quad F(h) = B(hn)$$

Gleichungen, die dem System (33. b) angehören, woraus nach (39) für $h = 1$

$$B(n) = \sum_k b(k)\, V(kn) \tag{41}$$

folgt. Hiermit sind die gesuchten Koeffizienten der in (32) angesetzten Entwicklung der Funktion $f(x)$ durch die Mittelwerte U und V ausgedrückt.

Der Bequemlichkeit halber sind nachstehend die Werte der Zahlen $a(k)$ und $b(k)$ für die ungeraden Zahlen von $k = 1$ bis $k = 99$ zusammengestellt. Es ist

$a(k) = b(k) = 0$ für

$k =$ 9, 25, 27, 45, 49, 63, 75, 81, 99;

$a(k) = +1$ für

$k =$ 1, 15, 21, 33, 35, 39, 51, 55, 57, 65,
69, 77, 85, 87, 91, 93, 95;

$a(k) = -1$ für

$k =$ 3, 5, 7, 11, 13, 17, 19, 23, 29, 31,
37, 41, 43, 47, 53, 59, 61, 67, 71, 73,
79, 83, 89, 97;

$b(k) = +1$ für

$k = 1, 3, 7, 11, 19, 21, 23, 31, 33, 43, 47, 57, 59, 65, 67, 69, 71, 77, 79, 83, 85, 93;$

$b(k) = -1$ für

$k = 5, 13, 15, 17, 29, 35, 37, 39, 41, 51, 53, 55, 61, 73, 87, 89, 91, 95, 97.$

Bei der Beweisführung war der Einfachheit halber vorausgesetzt worden, daß die $A(n)$ und $B(n)$ als die Glieder von unbedingt konvergenten Reihen angesehen werden können. Für die numerischen Anwendungen genügt jedoch die Voraussetzung, daß die trigonometrische Reihe für $f(x)$ gleichmäßig konvergiere. Zerlegt man nämlich in diesem Falle die Reihe in das Aggregat der Glieder bis zur Nummer n und in den überschießenden „Rest", so läßt sich für n ein fester Wert derart angeben, daß der Rest für alle x numerisch unmerklich ausfällt. Die Sache verhält sich also für die numerische Rechnung so, als ob $f(x)$ gleich der durch Fortlassung des Restes entstehenden endlichen Reihe wäre. Für solche endlichen Reihen bleiben aber die oben durchgeführten Schlüsse ohne weiteres gültig.

§ 97. Multipliziert man die Gleichungen (40) und (41) mit $\cos nx$ und $\sin nx$ und summiert darauf nach n, so erhält man die Darstellung

$$f(x) = A(0) + \sum_k \sum_n [a(k)\, U(kn) \cos nx + b(k)\, V(kn) \sin nx], \qquad (42)$$

die sich, wenn man nach den Größen U und V ordnet, in der Gestalt

$$f(x) = A(0) + \sum_m U(m)\, X_m + \sum_m V(m)\, Y_m, \quad (m = 1, 2, 3, \cdots)$$

schreiben läßt, wo die X_m, Y_m gewisse Verbindungen aus den Sinus und Kosinus der Vielfachen von x bedeuten. Führt man in (42) statt n den Quotienten $m:k$ ein, so erhält man

$$f(x) = A(0) + \sum_k \sum_m \left[a(k)\, U(m) \cos\left(\frac{mx}{k}\right) + b(k)\, V(m) \sin\left(\frac{mx}{k}\right) \right].$$

wobei natürlich die Summation nur über solche Wertepare k, m auszudehnen ist, für welche der Quotient $m:k$ gleich einer ganzen Zahl ist. Daraus folgen die Ausdrücke

$$X_m = \sum_k a(k) \cos\left(\frac{mx}{k}\right), \qquad Y_m = \sum_k b(k) \sin\left(\frac{mx}{k}\right), \qquad (43)$$

in denen man die Summation nach k jedesmal über die ungeraden Teiler von m, den Teiler $k = 1$ mit eingeschlossen, auszudehnen hat.

Führt man in den Ausdrücken X_m, Y_m statt x die unabhängige Veränderliche $y = \cos x$ oder $y = \sin x$ ein, so erhält man algebraische Formen von der Art, wie sie *Tschebyschef* aufgestellt hat, um eine Funktion $g(y)$, die in dem Intervall von $y = -1$ bis $y = +1$ gegeben ist, interpolatorisch darzustellen. Diese algebraische Einkleidung ist jedoch, wie schon oben erwähnt wurde, im allgemeinen eine dem Wesen der Sache fremde Künstelei. Man kann natürlich eine nichtperiodische Funktion $g(y)$ für das Intervall zwischen den Grenzen $y = \pm 1$ durch die Substitution $y = \cos x$ in eine periodische Funktion von x verwandeln und demgemäß entwickeln; man muß sich jedoch darüber klar sein, daß eine solche Umformung etwas überaus gequältes an sich hat, sobald nicht etwa die Werte $y = \pm 1$ zu den ausgezeichneten Stellen der Funktion $g(y)$ gehören.

§ 98. Zur Vervollständigung der Untersuchung ist noch zu erörtern, wie die Fehler der Funktionswerte in die Koeffizienten A, B eingehen. Die Größe $U(n)$ entsteht, wenn man vom Nullpunkte ausgehend den Umfang in $2n$ gleiche Teile teilt und die zugehörigen Funktionswerte mit alternierendem Vorzeichen mittelt. Demgemäß können wir schreiben

$$U(n) = \frac{1}{2n} \sum_p (-1)^p f\left(\frac{p\pi}{n}\right),$$

wo die Summation von $p = 0$ an über alle ganzzahligen p auszudehnen ist, für welche das Argument $p\pi : n$ noch unter 2π bleibt. Demgemäß ist nach (40)

$$A(n) = \sum_k \sum_p \frac{a(k)(-1)^p}{2kn} f\left(\frac{p\pi}{kn}\right), \tag{44}$$

wo k die ungeraden Zahlen $1, 3, 5, \cdot\cdot$ zu durchlaufen hat. Hierin vereinigen wir jetzt die Glieder, welche dasselbe Argument v enthalten, wobei v die Gestalt $q\pi : r$ besitzt, in der die Zahlen q und r als teilerfremd vorausgesetzt werden dürfen. Es ist dann

$$p\pi : kn = q\pi : r \quad \text{oder} \quad p = knq : r.$$

Nimmt der Bruch $n : r$, auf seine kleinste Benennung gebracht, die Gestalt $m : s$ an, so erhält man für p den Ausdruck $kmq : s$, der eine ganze Zahl sein muß. Da nun s nicht nur zu m, sondern als Teiler von r auch zu q teilerfremd ist, so muß die ungerade Zahl k durch s teilbar sein. Demgemäß kommen, da k ungerade ist, für s nur ungerade Zahlen und für k nur die Werte $s, 3s, 5s, \cdot\cdot$ in Betracht, woraus für p die Werte $mq, 3mq, \cdot\cdot$ folgen. Daher liefern

in (44) die Glieder, welche $f(v)$ enthalten, zu $A(n)$ den Beitrag

$$(-1)^{mq} f(v) \left[a(s) + \frac{a(3s)}{3} + \frac{a(5s)}{5} + \cdots \right] : 2ns. \tag{45}$$

Von dem Falle $s = 1$ abgesehen erfährt, weil dann s mindestens gleich 3 ist, der Fehler von $f(v)$ schon durch den Divisor $2ns$ eine erhebliche Verkleinerung. Für $s = 1$ erhält man als Faktor von $f(v)$ den Ausdruck

$$\pm \left[1 - \frac{1}{3} - \frac{1}{5} - \frac{1}{7} - \cdots \right] : 2n,$$

wo die einzelnen Terme in der Klammer den U Gliedern entsprechen, aus denen sich $A(n)$ zusammensetzt. Hiernach tritt also auch in dem Falle $s = 1$ eine erhebliche Verkleinerung des Fehlers von $f(v)$ ein.

Das vorstehende Verfahren läßt sich bei den Koeffizienten B wiederholen. Man hat zunächst

$$V(n) = \frac{1}{2n} \sum_p (-1)^p f\left(\frac{\pi}{2n} + \frac{p\pi}{n}\right),$$

$$B(n) = \sum_k \sum_p \frac{b(k)(-1)^p}{2kn} f\left(\frac{\pi}{2kn} + \frac{p\pi}{kn}\right).$$

Sucht man wieder die Glieder mit demselben Argument v zusammen und setzt $v = q\pi : r$, so wird

$$2p + 1 = 2knq : r$$

oder, wenn der Bruch $2n : r$, auf die kleinste Benennung gebracht, die Form $m : s$ annimmt,

$$2p + 1 = kmq : s.$$

Hierin muß wiederum k durch s teilbar sein, sodaß für k nur die Werte $k = ls$ in Betracht kommen, wo l die ungeraden Zahlen $1, 3, 5, \cdots$ zu durchlaufen hat. Demnach liefern die Glieder, welche $f(v)$ enthalten, zu $B(n)$ den Beitrag

$$\frac{f(v)}{2ns} \sum_l \frac{b(ls)(-1)^p}{l} = \frac{f(v)}{2ns} \sum_l \frac{a(ls)}{l} i^{ls - 1 + 2p}.$$

Den Exponenten von i können wir, weil $2p + 1$ jetzt gleich lmq wird, in der Gestalt

$$ls + lmq - 2 = l(s + mq) - 2$$

schreiben. Da nun s, m, q ungerade Zahlen sind, so ist $s + mq$ gleich einer geraden Zahl $2h$, sodaß die Potenz von i in der Gestalt

$$i^{2lh - 2} = -\left[(-1)^h\right]^l$$

erscheint, also für alle zulässigen ungeraden Zahlen l denselben Wert

$-(-1)^h$ oder $(-1)^{h+1}$ besitzt. Demnach nimmt der aus $f(v)$ entspringende Beitrag die Gestalt

$$(-1)^{h+1} f(v)\left[a(s)+\frac{a(3s)}{3}+\frac{a(5s)}{5}+\cdots\right]: 2ns$$

an, die sich mit den in (45) aufgestellten Ausdrucke deckt. Die Größen $B(n)$ verhalten sich also bezüglich der Fehlerwirkung nicht anders, als die Größen $A(n)$.

Da nach den vorstehenden Formeln die Fehlerverkleinerung zum Teil erst bei der Zusammensetzung der Größen U, V in Wirksamkeit tritt, so ist es angezeigt, in den Mitteln U, V zwei oder drei Decimalen mehr mitzunehmen, als in den Funktionswerten.

Achter Abschnitt.

Rekursionsformeln.

§ 99. Die numerische Bestimmung der Koeffizienten einer Reihenentwicklung gehört zu den Aufgaben, denen der Rechner bei den Anwendungen der Mathematik so häufig begegnet, daß sich für gewisse öfter wiederkehrende Formen ein, vorläufig allerdings noch kleiner, Bestand von besonderen Methoden herausgebildet hat. Hierher gehört außer der oben behandelten Entwicklung einer periodischen Funktion auch die Benutzung der sogenannten Rekursionsformeln, die jetzt besprochen werden soll.

Für die Funktion $f(x)$ sei die Reihenentwicklung

$$f(x) = A_0 g(x)_0 + A_1 g(x)_1 + \cdots$$

vorgeschrieben, in der die Koeffizienten A zu ermitteln sind, während die $g(x)$ bekannte und völlig bestimmte Funktionen von x bedeuten, deren Verhalten erst in Frage kommt, wenn die fertig gestellte Entwicklung weiter benutzt werden soll. Die analytische Untersuchung solcher Entwicklungen liefert nun wie bekannt oft den Nachweis, daß zwischen benachbarten Koeffizienten einfach gebaute Beziehungen bestehen, und es ist ohne weiteres klar, daß derartige „Rekursionsformeln" auch für die numerische Rechnung von Wichtigkeit werden können. In welcher Weise dies geschieht, hängt natürlich von den Umständen ab, sodaß man nicht daran denken kann, dafür ein festes überall passendes Schema aufstellen zu wollen; allgemein gültig ist nur die Forderung, daß bei dem rekursiven Verfahren die Sicherheit der Rechnung gewahrt bleibe. Es wird sich nun zeigen, daß dadurch häufig der einzuschlagende Weg ganz genau vorgeschrieben ist. Bei der Mannigfaltigkeit der Formen, die vorkommen können, muß ich mich darauf beschränken, bestimmte Typen, die seither in der Literatur häufiger aufgetreten sind, an passenden Beispielen zu erläutern.

Angenommen, es bestehe zwischen je drei aufeinander folgenden Koeffizienten eine Beziehung von der Gestalt

$$A_{n+1} = a_n A_n + b_n A_{n-1}, \qquad (1)$$

worin die a_n und b_n Größen bedeuten, die sich ohne besondere Mühe für jedes in Betracht kommende n berechnen lassen. Dann kann man, sobald A_0 und A_1 anderweitig ermittelt worden sind, aus (1) schrittweise die folgenden Koeffizienten $A_2, A_3, \cdots$ finden. Man übersieht auch leicht, daß ein solcher rückgreifender Rechnungsgang selbst dann von Vorteil sein kann, wenn sich jeder Koeffizient A_n bis zu einer gewissen Nummer hin ohne Schwierigkeit direkt aus seinem analytischen Ausdruck berechnen läßt, denn in der Regel wird der für die direkte Berechnung eines A_n erforderliche Rechnungsaufwand mit der Stellennummer n wachsen, wogegen bei der Benutzung von (1) der Arbeitsaufwand für den einzelnen Koeffizienten meistens merklich konstant bleibt. Man muß sich jedoch hüten, diese Bemerkung als ein allgemeines Gesetz aufzufassen, denn wenn in (1) von den Hilfsgrößen a, b eine beständig größer als Eins ist, so geht der Fehler, der einem A anhaftet, vergrößert in die folgenden Koeffizienten ein, und es läßt sich im voraus nicht ohne weiteres übersehen, ob eine solche steigende Fehlervergrößerung schließlich unschädlich bleibt oder aber das ganze Verfahren illusorisch macht. Im letzteren Falle müßte man sich darauf beschränken, die Formel (1) als Kontrolle für eine anderweite Berechnung der A-Größen zu benutzen, falls es nicht etwa gelingt, durch geeignete Umgestaltungen die Schwierigkeit zu beseitigen. Die nachstehenden Beispiele können eine Vorstellung davon geben, wie man sich unter Umständen zu helfen vermag.

§ 100. Als erstes Beispiel sollen die sogenannten Kugelfunktionen dienen. Dieselben entstehen, wenn man den Ausdruck

$$W = 1 : \sqrt{(1 - 2xt + t^2)} \tag{2}$$

nach Potenzen von t in die Reihe

$$W = P_0 + P_1 t + P_2 t^2 + \cdots = \sum_n P_n t^n \tag{3}$$

entwickelt, deren Ableitung W' die Gestalt

$$W' = P_1 + 2P_2 t + \cdots = \sum_n n P_n t^{n-1} \tag{4}$$

annimmt. Setzt man zunächst $y = 2xt - t^2$, so kann man die binomische Entwicklung

$$W = 1 : \sqrt{(1 - y)} = e_0 + e_1 y + e_2 y^2 + \cdots \tag{5}$$

herstellen, wo

$$e_0 = 1, \quad e_1 = \frac{1}{2} \quad e_2 = \frac{1 \cdot 3}{2 \cdot 4}, \cdots$$

ist. Der allgemeine Ausdruck für e_n läßt sich in der Gestalt

$$e_n = (2n)! : (2^n n!)^2 \tag{6.a}$$

schreiben, woraus noch

$$e_n : e_{n-1} = (2n-1) : 2n \tag{6.b}$$

folgt. Entwickelt man in (5) die Potenzen von y nach t und bezeichnet mit $(n)_p$ wie früher den p^{ten} Binomialkoeffizienten für die Potenz n, so wird

$$P_n = e_n(2x)^n - e_{n-1}(n-1)_1(2x)^{n-2} + e_{n-2}(n-2)_2(2x)^{n-4} - \cdots, \tag{7}$$

woraus z. B. für die niedrigsten n

$$P_0 = 1, \quad P_1 = x, \quad 2P_2 = 3x^2 - 1$$

folgt. Die Kugelfunktion P_n erscheint also als ein Polynom n^{ter} Ordnung von x, das mit n gleichzeitig gerade oder ungerade ist, sodaß man sich bei der weiteren Untersuchung auf den Fall positiver x beschränken darf.

Um eine Rekursionsformel zu gewinnen schreibt man (2) in der Gestalt

$$W^2(1 - 2xt + t^2) = 1,$$

woraus durch Differentiation nach t die Beziehung

$$W'(1 - 2xt + t^2) + W(t - x) = 0$$

folgt. Führt man hierin statt W und W' die in (3) und (4) aufgestellten Reihen ein, so müssen sich die Glieder, welche dieselbe Potenz von t enthalten, gegenseitig aufheben. Daraus fließt die gesuchte Formel in der Gestalt

$$0 = (n+1)P_{n+1} - (2n+1)xP_n + nP_{n-1},$$

die mit den Abkürzungen

$$a_n = (2n+1)x : (n+1), \quad b_n = n : (n+1) \tag{8.a}$$

in die Formel

$$P_{n+1} = a_n P_n - b_n P_{n-1} \tag{8.b}$$

übergeht. Die Hilfsgröße b_n ist beständig kleiner als Eins, und das Gleiche gilt von a_n, solange $2x$ ein echter Bruch ist. In diesem Falle besteht also kein Hindernis die Rekursionsformel unmittelbar in der Gestalt (8) anzuwenden, um von P_0 und P_1 ausgehend die übrigen P zu berechnen. Sind dagegen die a_n größer als Eins, so wird die Rechnung nach (8) vorläufig unsicher. Um nun das Verhalten der P besser zu übersehen, werde für den Augenblick x als sehr groß angenommen, sodaß in der Reihe (7) das Anfangsglied $e_n(2x)^n$ als das Hauptglied erscheint. Setzt man daraufhin

$$P_n = e_n(2x)^n Q_n,$$

so kann man mit einer kleinen Reduktion statt der Gleichungen (8) die Beziehungen

$$Q_{n+1} = Q_n - c_n Q_{n-1}, \qquad c_n = n^2 : [(4n^2 - 1)x^2] \tag{9}$$

ansetzen, deren Durchrechnung mit den Werten $Q_0 = 1$, $Q_1 = 1$ zu beginnen hat.

Die Benutzung von (9) ist völlig sicher, solange x oberhalb 0.5 liegt. Da Q_n nachher mit $e_n(2x)^n$ zu multiplizieren ist, so liegt es nahe, sogleich auf die Logarithmen der Q-Größen auszugehen, also die rechte Seite der ersten Gleichung (9) jedesmal mit Additions- oder Subtraktionslogarithmen zu bilden. In diesem Falle liegt aber keine Nötigung mehr vor, erst die Q-Größen einzuführen, denn man kann dann ebensogut auch sogleich nach (8) rechnen. Diese Rechnung wird noch übersichtlicher, wenn man statt der Größen Log P zunächst die Logarithmen der P-Quotienten aufsucht. Setzt man

$$R_n = P_n : P_{n-1},$$

so ist $R_1 = P_1 : P_0 = x$, ferner wird der Reihe nach

$$P_1 = R_1, \qquad P_2 = P_1 R_2, \qquad P_3 = P_2 R_3, \quad \text{u. s. w.},$$

während die R-Größen durch die Gleichungen

$$R_{n+1} = a_n - \frac{b_n}{R_n}$$

zusammenhängen.

Als Beispiel diene die Annahme Log x = 9.76148. Hierzu erhält man zunächst unter Benutzung fünfstelliger Tafeln

n	a	Log a	b	Log b
1	$3x:2$	9.93757	1 : 2	9.69897
2	$5x:3$	9.98333	2 : 3	9.82391
3	$7x:4$	0.00452	3 : 4	9.87506

Daraus ergeben sich von $R_1 = x$ aus die nachstehenden, sogleich zu erläuternden Zahlen:

n	Log R	Log P	P
1	9.76148	9.76148	+ 0.57740
2	6.15249	5.91397	+ 0.00008
3	3.67133 n	9.58530 n	− 0.38485
4	0.00459	9.58989 n	− 0.38895

Beim Aufschlagen von Log R_2 ergibt sich, daß R_2 sehr klein ist und

daß der Logarithmus nur sehr unsicher gefunden wird: er ist zwischen 6.17749 und 6.12749 zu suchen. Vorstehend ist das Mittel dieser beiden Grenzen angesetzt. Ferner wird R_3 sehr groß und wegen der Unsicherheit von Log R_2 ebenfalls nur unsicher gefunden. Diese beiden Unsicherheiten heben sich jedoch in dem Produkt $R_2 R_3$ gegenseitig auf; in der Tat stimmen die oben angesetzten Werte der P-Größen bis auf eine Einheit der letzten Stelle mit den wahren Werten überein.

Wendet man für den Kettenbruch

$$a + \cfrac{b}{c + \cfrac{d}{e + \cdots}}$$

die gedrängtere Schreibweise

$$a + b : c + d : e + \cdots$$

an, so beruht die Ermittlung der R-Größen offenbar darauf, daß man den Kettenbruch

$$R_{n+1} = a_n - b_n : a_{n-1} - b_{n-1} : a_{n-2} - \cdots,$$

der bei a_0 abbricht, von unten herauf berechnet.

§ 101. Als zweites Beispiel diene die Entwicklung des Ausdruckes

$$W = 1 : \sqrt{(1 - 2x \cos y + x^2)} \tag{10}$$

in die trigonometrische Reihe

$$W = A_0 + 2A_1 \cos y + 2A_2 \cos 2y + \cdots. \tag{11}$$

Setzt man $\exp(iy) = z$ und zerlegt das Trinom in seine beiden Faktoren, so wird

$$W = (1 - xz)^{-\frac{1}{2}} (1 - xz^{-1})^{-\frac{1}{2}}.$$

Entwickelt man

$$(1 - xz)^{-\frac{1}{2}} = c_0 + c_1 xz + c_2 x^2 z^2 + \cdots,$$

$$(1 - xz^{-1})^{-\frac{1}{2}} = c_0 + c_1 xz^{-1} + c_2 x^2 z^{-2} + \cdots,$$

so erhält man durch Ausmultiplizieren der beiden Reihen für die Koeffizienten A_n sofort die Darstellung

$$A_n = c_0 c_n x^n + c_1 c_{n+1} x^{n+2} + \cdots. \tag{12}$$

Läßt man in A_n für den Index n auch negative Werte zu, mit der Festsetzung, daß $A_{-n} = A_n$ sein solle, so kann man die Reihe für W und für die nach y genommene Ableitung W' in der Gestalt

$$W = \sum_n A_n z^n, \qquad W' = \sum_n i n A_n z^n \tag{13}$$

schreiben, wo die Summation nach n von $-\infty$ bis $+\infty$ läuft. Schreibt

man ferner (10) in der Gestalt

$$(1 - 2x\cos y + x^2)W^2 = 1$$

und nimmt davon die Ableitung nach y, so wird

$$2(1 - 2x\cos y + x^2)W' + 2x\sin y\, W = 0.$$

Hieraus ergibt sich durch Einsetzung der in (13) aufgestellten Reihen

$$\sum_n A_n[2in(1 + x^2)z^n - 2inx(z + z^{-1})z^n - ix(z - z^{-1})z^n] = 0.$$

Spaltet man nunmehr nach den Potenzen von z, so entsteht die Rekursionsformel

$$(2n + 1)xA_{n+1} - 2n(1 + x^2)A_n + (2n - 1)xA_{n-1} = 0, \qquad (14)$$

die zunächst wieder in der Gestalt

$$A_{n+1} = a_n A_n - b_n A_{n-1}, \qquad (15.\text{a})$$

$$a_n = [2n(1 + x^2)] : [(2n + 1)x], \quad b_n = (2n - 1) : (2n + 1) \qquad (15.\text{b})$$

geschrieben werden soll.

Versucht man, von den Größen A_0 und A_1 ausgehend, nach (15) zu rechnen, so stellt sich heraus, daß man nicht zum Ziele kommt. Am raschesten läßt sich das übersehen, wenn man x sehr klein nimmt. Denn in diesem Falle wird a_n sehr groß, sodaß der Fehler jedes A_n stark vergrößert in das nächste A_{n+1} übergeht. Setzt man ferner, um die Rechnung rein logarithmisch durchzuführen, die Quotienten

$$B_n = A_n : A_{n-1}$$

an und schreibt demgemäß (15.a) in der Gestalt

$$B_{n+1} = a_n - \frac{b_n}{B_n}, \qquad (16)$$

so ist die linke Seite nach x von der Ordnung $+1$, während die beiden Terme rechter Hand von der Ordnung -1 sind. Es tritt also fortwährend der nachteilige Umstand ein, daß eine kleine Zahl als Differenz aus zwei großen gefunden wird. Diese Schwierigkeiten lassen sich nun in einfacher Weise durch einen zuerst von *Hansen* ausgenutzten Kunstgriff umgehen, der darin besteht, daß man die Quotienten B nicht in der Reihenfolge $B_1, B_2, \cdots$, sondern in umgekehrter Ordnung rechnet. Um die hierbei auftretenden Beziehungen zu übersehen, hat man zunächst den Zusammenhang zwischen den Fehlern der B-Größen klar zu stellen.

§ 102. Die in (11) und (12) aufgestellten Reihen setzen voraus, daß x ein echter Bruch sei. Für die numerische Rechnung darf überdies x nicht zu nahe bei Eins liegen, da andernfalls die Reihe (11)

viel zu langsam konvergiert; beispielsweise erreicht für $x = 0.9$ das Anfangsglied in A_{10} noch den Betrag 0.06.

Setzt man nun in (12) $A_n = c_0 c_n x^n C_n$, so erhält man für die C-Größen die Reihe

$$C_n = 1 + \frac{1}{2}\frac{2n+1}{2n+2}x^2 + \frac{1\cdot 3}{2\cdot 4}\frac{(2n+1)(2n+3)}{(2n+2)(2n+4)}x^4 + \cdots,$$

die mit wachsendem n gegen die Reihe

$$1 + \frac{1}{2}x^2 + \frac{1\cdot 3}{2\cdot 4}x^4 + \cdots$$

konvergiert, sodaß der Quotient aus zwei benachbarten C-Größen den Grenzwert Eins besitzt. Multipliziert man die Reihe für C_n mit dem Quotienten $(2n-1):2n$, so sind die einzelnen Glieder der neuen Reihe kleiner als die entsprechenden Glieder gleicher Nummer in der Reihe für C_{n-1}, sodaß $(2n-1)C_n$ stets kleiner als $2nC_{n-1}$ ist. Da nun

$$\begin{aligned} B_n = A_n : A_{n-1} &= (c_n x C_n) : (c_{n-1} C_{n-1}) \\ &= [(2n-1)xC_n] : [2nC_{n-1}] \end{aligned}$$

wird, so ist B_n stets kleiner als x und konvergiert mit wachsendem n von unten her gegen x als Grenzwert. Dies festgestellt schreiben wir die Gleichung (16) in der Gestalt

$$B_n B_{n-1} - a_n B_n + b_n = 0 \tag{17}$$

und benutzen sie, um jetzt von einem genäherten B_{p-1} ausgehend rückwärts die Größen $B_p, B_{p-1} \cdots$ zu berechnen. Um hierbei die Fortpflanzung des Fehlers zu übersehen, bezeichnen wir die richtigen Werte der B-Größen wie bisher durch den Buchstaben B und setzen die fehlerhaften Werte in der Gestalt

$$B(1-f) \quad \text{oder} \quad B:(1-g)$$

an, wo dann

$$(1-f)(1+g) = 1, \quad f = g:(1+g), \quad g = f:(1-f)$$

ist. Sind die Fehler klein, so sind f und g merklich einander gleich und liefern, mit dem Modulus der gemeinen Logarithmen multipliziert, die Verbesserung des bei der Rechnung benutzten Wertes von Log B, während das Produkt Bf oder Bg die Verbesserung des Numerus angibt. Ersetzt man nun in (17) die Größen B_n und B_{n-1} durch die zusammengehörigen fehlerhaften Werte

$$B_n:(1+g_n) \quad \text{und} \quad B_{n-1}(1-f_{n-1}),$$

so erhält man die Gleichung

$$B_n B_{n-1}(1-f_{n-1}) - a_n B_n - b_n(1+g_n) = 0.$$

die mit (17) verbunden zu den Beziehungen

$$g_n : f_{n+1} = B_n B_{n+1} : b_n = B_{n+1} : (a_n - B_{n+1})$$

führt. Setzt man hierin die für große n genäherten Werte

$$B_{n+1} = x, \quad a_n = (1 + x^2) : x$$

ein, so wird

$$g_n = x^2 f_{n+1}, \tag{18}$$

woraus zu entnehmen ist, daß bei der betrachteten Rückwärtsrechnung eine beständige Verkleinerung des im Ausgangswerte begangenen Fehlers und zwar in geometrischer Progression stattfindet.

Setzt man die aus (17) folgende Gleichung

$$B_n = b_n : (a_n - B_{n+1}) \tag{19}$$

für die Nummern $n, n + 1, \cdot\cdot$ an und eliminiert die höheren B, so entsteht daraus der unendliche Kettenbruch

$$B_n = b_n : a_n - b_{n+1} : a_{n+1} - b_{n+2} : a_{n+2} - \cdot\cdot, \tag{20}$$

der nach den vorstehenden Darlegungen sicher konvergiert.

§ 103. An der Hand der gefundenen Formeln ergibt sich nunmehr folgender Rechnungsgang. Da

$$A_1 = A_0 B_1, \quad A_2 = A_1 B_2, \quad A_3 = A_2 B_3, \cdot\cdot \tag{21}$$

oder allgemein

$$A_n = A_0 B_1 B_2 \cdot\cdot B_n$$

ist, so wird, weil im vorliegenden Falle die B-Größen von unten her gegen x konvergieren, A_n kleiner als $A_0 x^n$. Auf Grund dieses Umstandes kann man mit einem rohen Überschlagswerte von A_0 zunächst abschätzen, bis zu welcher Nummer hin die gesuchten Koeffizienten A in Betracht kommen. Ist A_p der letzte Koeffizient, der zu berücksichtigen ist, so wird Log B_p gebraucht, wenn auch nur mit einer oder höchstens zwei Decimalen. Man könnte nun den Ausgangswert B_p nach (20) bilden, indem man zu dem angegebenen Kettenbruch von oben her nach den bekannten Vorschriften die nötige Anzahl von Näherungsbrüchen rechnet. Da jedoch hierbei die Zähler und Nenner der einzelnen Näherungsbrüche getrennt gebildet werden müssen, obgleich es nur auf die Werte der Brüche selber ankommt, so ist es in der Regel bequemer, nur nach (19) zu rechnen, dabei aber die Rechnung mit der Näherung $B_m = x$ zu beginnen, wo die Nummer m um eine passende Anzahl von Einheiten oberhalb p genommen wird. Sind die B-Größen gefunden, so erhält man die Koeffizienten A nach (21), sobald A_0 aus der Reihe (12) oder auf irgend einem anderen Wege ermittelt ist.

Der beschriebene Weg gewährt offenbar noch den Vorteil, daß man nur das Anfangsglied A_0 besonders zu berechnen hat, während die gewöhnliche Benutzungsart der Rekursionsformel (15) voraussetzt, daß außer A_0 auch A_1 bekannt sei.

Im vorliegenden Falle, wo es sich um die nach y periodische Funktion

$$W = 1 : \sqrt{(1 - 2x\cos y + x^2)}$$

handelt, kann man übrigens A_0 sehr bequem auf folgende Weise herleiten. Man teilt den Kreisumfang für y vom Nullpunkte aus in q gleiche Teile, berechnet für die Teilpunkte die Werte von W und bildet daraus das arithmetische Mittel M. Dann ist nach den Sätzen über die trigonometrischen Reihen

$$M = A_0 + 2A_q + 2A_{2q} + \cdots = A_0\left(1 + \frac{2A_q}{A_0} + \frac{2A_{2q}}{A_0} + \cdots\right),$$

wo die Quotienten in der Klammer durch die B Größen gegeben sind, sodaß man aus M unmittelbar auch A_0 findet.

§ 104. Sieht man von den Besonderheiten des vorstehenden Beispiels ab, so erkennt man, daß der *Hansen*sche Kunstgriff sich mit Vorteil anwenden läßt, sobald erstens in der Rekursionsformel

$$A_{n+1} = a_n A_n - b_n A_{n-1}$$

die Größen a, b auf einen konvergenten Kettenbruch von der Gestalt

$$b_n : a_n - b_{n+1} : a_{n+1} - \cdots$$

führen und zweitens $\operatorname{Log} A_0$ mit der genügenden Genauigkeit bekannt ist. Daß die zweite Bedingung besonders hervorgehoben werden muß, läßt sich am besten an einem Beispiel nachweisen.

Definiert man die *Bessel*sche Funktion $J(x)_n$ für ganzzahlige n durch die Reihe

$$J(x)_n = \frac{1}{0!\,n!}\left(\tfrac{1}{2}x\right)^n - \frac{1}{1!\,(n+1)!}\left(\tfrac{1}{2}x\right)^{n+2} + \frac{1}{2!\,(n+2)!}\left(\tfrac{1}{2}x\right)^{n+4} - \cdots,$$

so besteht die an der Hand der vorstehenden Reihe leicht zu beweisende Beziehung

$$xJ_{n+1} - 2nJ_n + xJ_{n-1} = 0,$$

die in der bisherigen Bezeichnungsweise auf die Größen

$$a_n = 2n : x, \quad b_n = 1$$

führt. Der zugehörige Kettenbruch konvergiert, wenn auch nicht sogleich am Anfange, so doch später sehr rasch, sodaß die Quotienten $J_n : J_{n-1}$ ohne Schwierigkeit ermittelt werden können. Wenn nun aber, wie das für gewisse Werte von x eintreten kann, J_0 verschwindet, so

wird der Quotient $J_1 : J_0$ unendlich, d. h. sämtliche J außer J_0 erscheinen in der unbestimmten Form $0 \cdot \infty$. Ferner tritt ein solches Versagen der Rechnungsvorschrift für die numerische Ausführung schon dann ein, wenn J_0 nicht null, sondern nur sehr klein ist. Man ist dann genötigt, in der Reihe der J rückwärts nicht bis J_0, sondern nur bis J_1 zu gehen, und außer J_0 auch J_1 direkt aus der Reihenentwicklung oder sonstwie zu bestimmen.

Die vorstehenden Fälle mögen hier genügen. Im übrigen versteht es sich von selbst, daß man bei der Verwertung einer Rekursionsgleichung sowohl die Gestalt dieser Gleichung, als auch die Natur der darin auftretenden Funktionen zu beachten hat, um jedesmal den vorteilhaftesten Rechnungsgang ausfindig zu machen.

Neunter Abschnitt.

Interpolation im weiteren Sinne.

§ 105. Bei der Behandlung der Interpolation im zweiten Abschnitte war bereits angedeutet worden, daß die dort untersuchte Aufgabe als Teil eines allgemeineren Problems, nämlich der Interpolation im weiteren Sinne des Wortes, anzusehen sei. Der bei der früheren Aufgabe befolgte Gedankengang ließ sich, geometrisch gesprochen, folgendermaßen formulieren. Auf einer Kurve $y = f(x)$ ist eine Reihe von Punkten, die zu äquidistanten Ordinaten gehören, gegeben und zwar für das ganze Abscissengebiet, das bei der Benutzung der Kurve in Betracht kommen kann. Um einen neuen Kurvenpunkt einzuschalten, wird ein Bogenstück herausgegriffen, das einige wenige Intervalle umfaßt und den gesuchten Punkt einschließt. Weiter wird durch die gegebenen Punkte des Bogenstückes eine Parabel höherer Ordnung gelegt und diese als Ersatz an Stelle der vorgelegten Kurve benutzt. Damit ein solcher Ersatzausdruck, der die Bedeutung einer für $f(x)$ substituierten Näherungsformel besitzt, für den Rechner brauchbar sei, war zweierlei zu fordern: erstens mußte die benutzte Formel hinreichend einfach sein, zweitens mußte an der betrachteten Stelle ein hinreichend enger „*Anschluß*“ zwischen der vorgelegten Kurve und der dafür substituierten Parabel stattfinden. Beide Forderungen ließen sich, wie wir gesehen haben, gleichzeitig erfüllen, wenn die Funktion $f(x)$ für das betrachtete Bogenstück durch die *Taylor*sche Reihe darstellbar ist, und wenn ferner nur wenige Glieder dieser Reihe für die Rechnung zu berücksichtigen sind.

Man kann sich nun vornehmen, die damals behandelte Aufgabe zu erweitern. Zunächst kann man nämlich, indem man vorläufig die Beschränkung auf ein kurzes Bogenstück festhält, die Voraussetzung fallen lassen, daß die Ordinaten äquidistant seien. In diesem Falle, der dem Rechner übrigens verhältnismäßig selten entgegentritt, ist die Lösung der Aufgabe wiederum durch die *Lagrange*sche Interpolationsformel gegeben: der ganze Unterschied gegen die früher entwickelten Rechnungsvorschriften besteht darin, daß nunmehr die Vereinfachungen fortfallen, die bei äquidistanten Ordinaten eintreten. Ebenso bleiben die vorhin angeführten auf die *Taylor*sche Reihe bezüglichen Voraus-

setzungen für die rechnerische Brauchbarkeit der Formel bestehen. Aus diesem Grunde ist es nicht nötig, hier auf den Fall ungleicher Intervalle noch besonders einzugehen.

Eine andere Erweiterung hingegen und zwar die weitaus wichtigere geht darauf aus, die Beschränkung auf ein kurzes Bogenstück zu beseitigen. Es handelt sich dann darum, die vorgelegte Kurve $y = f(x)$ durch eine andere mit der Gleichung $y = F(x)$ zu ersetzen, die folgenden Bedingungen genügt: erstens muß der Ausdruck von $F(x)$ für die numerische Rechnung genügend bequem ausfallen, zweitens muß durchweg ein hinreichend enger Anschluß zwischen den beiden Kurven bestehen, sodaß die Differenzen $f(x) - F(x)$ für das ganze in Betracht kommende Abscissengebiet als unerheblich anzusehen sind.

Der Anlaß, eine solche „*interpolatorische Darstellung*" für die vorgelegte Kurve zu substituieren, kann durch sehr verschiedene Umstände gegeben sein. Es kann vorkommen, daß die analytischen Beziehungen, aus denen $f(x)$ entstanden ist, vollkommen offen liegen, daß jedoch die Gestalt von $f(x)$ äußerst verwickelt ausfällt: in diesem Falle dient die Darstellung durch $F(x)$ zur Entlastung des Rechners. Andererseits kann es aber auch vorkommen, daß das analytische Bildungsgesetz von $f(x)$ nur teilweise oder auch garnicht bekannt ist: dann verfolgt die Aufsuchung von $F(x)$ den Zweck, an die Stelle der numerischen Tabelle, in der die Koordinatenpaare (xy) für die gegebenen Kurvenpunkte zusammengestellt sind, einen analytischen Ausdruck zu setzen, der den Inhalt jener Tabelle zu „vertreten" geeignet ist.

§ 106. Versucht man die vorstehend beschriebene Aufgabe für einen bestimmten Fall zu lösen, so stößt man jedesmal zuerst auf die Frage, wie der Ausdruck $F(x)$ zu wählen sei. Daß man hierbei nach Möglichkeit alle Anhaltspunkte benutzen wird, die bezüglich der darzustellenden Funktion $f(x)$ bekannt sind, versteht sich von selbst; es ist jedoch klar, daß man $F(x)$ probeweise auf gut Glück anzusetzen hat, wenn die analytische Beschaffenheit von $f(x)$ nur ungenügend bekannt ist. Der hierbei mit Vorliebe eingeschlagene Weg besteht nun darin, daß man für $F(x)$ eine Reihe von der Form

$$F(x) = X_1 g(x)_1 + X_2 g(x)_2 + \cdots \qquad (1)$$

aufstellt, in der die Funktionen $g(x)$, von der Veränderlichen x abgesehen, numerisch vollständig bestimmt sind, wogegen die Koeffizienten X gewisse vorläufig verfügbare Konstanten bedeuten, die so anzusetzen sind, daß ein möglichst enger Anschluß von $F(x)$ an $f(x)$ entsteht. Einfache und häufig vorkommende Beispiele für die Form (1) bietet der Fall der gewöhnlichen Potenzreihe und der früher untersuchten trigonometrischen Reihe.

Der Umstand, daß die sogenannten „*Parameter*“ der gesuchten Darstellung, nämlich die unbekannten X, in (1) linear auftreten, bildet bei der Ermittelung der Parameter eine wesentliche Erleichterung; dieser Vorteil ist so erheblich, daß man die lineare Form auch da herbeizuführen sucht, wo sie aus irgend welchen Gründen ursprünglich nicht vorhanden ist. Der Kunstgriff, den man hierbei anwendet, ist folgender. Wenn man durch bestimmte Umstände veranlaßt worden ist, für $F(x)$ vorerst einen in den Parametern nichtlinearen Ausdruck $G(x, X_1, X_2, \cdots)$ aufzustellen, so denke man sich für die X_p zunächst genäherte „*vorläufige*“ Werte Y_p ermittelt und darauf $X_p = Y_p + Z_p$ gesetzt, wo die „*Verbesserungen*“ Z_p nunmehr als die Unbekannten des Problems auftreten. Sind nun die für die Z_p zu erwartenden Werte genügend klein, so kann man nach der *Taylor*schen Reihe unter Beschränkung auf die Glieder erster Ordnung den Ausdruck

$$F(x) = G(x, Y_1, Y_2, \cdots) + G_1 Z_1 + G_2 Z_2 + \cdots$$

ansetzen, wo die Koeffizienten $G_1, G_2, \cdots$ die partiellen Ableitungen des ersten Gliedes der rechten Seite nach $Y_1, Y_2, \cdots$ bedeuten. Damit ist offenbar die lineare Form für $F(x)$ hergestellt.

Kehrt man wieder zu der Form (1) zurück, so liefert jeder gegebene Punkt der Kurve $y = f(x)$ eine Gleichung von der Gestalt

$$f(x) = X_1 g(x)_1 + X_2 g(x)_2 + \cdots, \tag{2}$$

in der bis auf die Parameter X alles numerisch bekannt ist. Man könnte also, wenn die gewählte Reihenform bezüglich der Gliederanzahl keiner Beschränkung unterworfen ist, im allgemeinen in sämtlichen gegebenen Kurvenpunkten eine vollständige Übereinstimmung zwischen $f(x)$ und $F(x)$ herbeiführen, indem man einfach die Reihe mit ebensovielen Gliedern ansetzt, als Kurvenpunkte gegeben sind. Ein solches Vorgehen wäre jedoch im allgemeinen zweckwidrig, und zwar aus folgenden Gründen. Zunächst bliebe die Forderung hinreichender Einfachheit unerfüllt, sobald die Menge der gegebenen Kurvenpunkte einen irgendwie erheblichen Betrag annimmt: wenn fünfzig Punkte gegeben sind, so würde die dazu gehörige, aus fünfzig Gliedern der Reihe (1) bestehende Formel der Regel nach als für den ernstlichen Gebrauch ungeeignet angesehen werden. Sodann bietet aber auch der für die sämtlichen gegebenen Punkte erzwungene genaue Anschluß keinerlei Gewähr dafür, daß auch in den Zwischenstellen eine befriedigende Übereinstimmung zwischen $f(x)$ und $F(x)$ stattfinde. Denn das Verhalten der Differenzen $f - F$ in den Zwischenstellen hängt wesentlich von den Eigenschaften der Funktionen $f(x)$ und $g(x)$ ab, während der genaue Anschluß in den gegebenen Kurvenpunkten ohne Rücksicht auf jene Eigenschaften einzig und allein

dadurch bewirkt worden ist, daß man die Menge der verfügbaren Parameter gleich der Menge der gegebenen Punkte genommen hat.

Läßt man mit Rücksicht auf die vorstehenden Bemerkungen nunmehr als Regel zu, daß die Anzahl als Parameter kleiner als die Anzahl der gegebenen Kurvenpunkte ist, so treten in dem Gleichungssystem (2) mehr Gleichungen als Unbekannte auf. Ist nun n die Anzahl der Gleichungen, p die der verfügbaren Parameter, so kann man zwar immer noch den genauen Anschluß zwischen $f(x)$ und $F(x)$ in p von den gegebenen Punkten erzwingen, wird es jedoch im allgemeinen vorziehen, die n Differenzen $f - F$ in ihrer Gesamtheit möglichst herunterzudrücken, statt einige von ihnen genau gleich Null zu machen. Damit wird man zu einer Aufgabe geführt, die zugleich das Grundproblem der sogenannten Ausgleichsrechnung bildet, und die nunmehr mit etwas abgeänderten Bezeichnungen folgendermaßen formuliert werden soll: Für die p Unbekannten $X_1, X_2, \cdots$ sind Gleichungen von der Form

$$\left.\begin{aligned} n_1 &= a_1 X_1 + b_1 X_2 + c_1 X_3 + \cdots, \\ n_2 &= a_2 X_1 + b_2 X_2 + c_2 X_3 + \cdots, \\ &\text{u. s. w.} \end{aligned}\right\} \quad (3)$$

angesetzt worden, deren Anzahl m größer als p ist; die Größen $n_1, n_2, \cdots$ sind so zu beschaffen, daß die angesetzten Gleichungen durch kein Wertsystem der Unbekannten gleichzeitig befriedigt werden können; infolgedessen führt jedes System der Unbekannten auf *Widersprüche* von der Form

$$\left.\begin{aligned} D_1 &= n_1 - a_1 X_1 - b_1 X_2 - c_1 X_3 - \cdots, \\ D_2 &= n_2 - a_2 X_1 - b_2 X_2 - c_2 X_3 - \cdots, \\ &\text{u. s. w.,} \end{aligned}\right\} \quad (4)$$

die nicht gleichzeitig zum Verschwinden gebracht werden können; gesucht wird ein System der X oder, wie wir kurz sagen wollen, eine „*Lösung*", bei der die Gesamtheit der Widersprüche auf einen möglichst niedrigen Betrag herabgedrückt ist.

Es ist nun zu untersuchen, ob und wie weit sich die gestellte Aufgabe in bestimmter und einwandfreier Weise lösen läßt.

§ 107. Bedeuten L und L' zwei Lösungen, deren zugehörige Widersprüche mit D und D' bezeichnet werden mögen, so wird, falls jedes D numerisch kleiner als das entsprechende D' ist, die Lösung L den Vorzug verdienen. Denn die Widersprüche sind in unserem Falle nichts anderes, als die Abweichungen der vorgelegten Funktion $f(x)$ von ihrer interpolatorischen Darstellung $F(x)$, sodaß bei der Lösung L der Anschluß durchweg enger ausfällt als bei L'. Sind dagegen die D ihrem numerischen Betrage nach teils kleiner, teils größer als die ent-

sprechenden D', so läßt sich vorläufig gar keine Entscheidung treffen, ob L oder L' vorzuziehen sei. Vielmehr muß in einem solchen Falle irgend eine weitere Bestimmung hinzugefügt werden, durch die man die vorhandene Ungewißheit aufheben kann. Unter den Vorschlägen, die man in dieser Beziehung gemacht hat, sind seither für den ernsthaften Gebrauch nur zwei in Betracht gekommen: der eine führt zu der sogenannten „*Methode der kleinsten Quadrate*", der andere zu dem *Interpolationsverfahren von Cauchy.*

Die Methode der kleinsten Quadrate beruht auf der Festsetzung, daß unter den verschiedenen Lösungen, die *vorläufig* als gleich annehmbar anzusehen sind, diejenige als die „annehmbarste" gelten solle, für welche die Quadratsumme der Widersprüche kleiner ausfällt, als für jede andere Lösung. Bezeichnet man der Kürze halber die Produktsumme

$$P_1 Q_1 + P_2 Q_2 + \cdots + P_m Q_m$$

durch das Symbol $[PQ]$, so erscheint die vorstehende Festsetzung in der Gestalt

$$[DD] = \text{Minimum}.$$

Hiernach hat man, um die Bedingungen für die gesuchte Lösung aufzustellen, von dem Ausdrucke $[DD]$ nach den einzelnen Unbekannten die partiellen Ableitungen erster Ordnung zu bilden und gleich Null zu setzen. Dies liefert die Gleichungen

$$[Da] = 0, \quad [Db] = 0, \quad [Dc] = 0, \quad \text{u. s. w.},$$

die ausgeschrieben zu dem System

$$\left.\begin{aligned} [na] &= [aa]X_1 + [ab]X_2 + \cdots, \\ [nb] &= [ba]X_1 + [bb]X_2 + \cdots, \\ [nc] &= [ca]X_1 + [cb]X_2 + \cdots, \\ &\text{u. s. w.} \end{aligned}\right\} \qquad (5)$$

führen. Um das vorstehende System, das die sogenannten „*Normalgleichungen*" der Aufgabe enthält, aufzulösen, verfährt man in folgender Weise. Man multipliziert die erste Gleichung mit dem Quotienten $[ba]:[aa]$ und zieht die solchergestalt abgeänderte Gleichung von der zweiten Gleichung des Systems ab, wobei die Unbekannte X_1 herausfällt. Ferner multipliziert man die erste Gleichung mit dem Quotienten $[ca]:[aa]$ und zieht von der dritten Gleichung ab, u. s. w., bis das ganze System erschöpft ist. Hierdurch entsteht ein von X_1 freies reduziertes System, das wir in der Gestalt

$$\begin{aligned} [nb\cdot 1] &= [bb\cdot 1]X_2 + [bc\cdot 1]X_3 + \cdots, \\ [nc\cdot 1] &= [cb\cdot 1]X_2 + [cc\cdot 1]X_3 + \cdots, \\ &\text{u. s. w.} \end{aligned}$$

schreiben. Dieses System wird wieder in derselben Weise behandelt, indem man die erste Gleichung benutzt, um X_2 aus den übrigen Gleichungen herauszuschaffen, sodaß ein neues reduziertes System von der Gestalt

$$[nc\cdot 2] = [cc\cdot 2]X_3 + [cd\cdot 2]X_4 + \cdots,$$
$$[nd\cdot 2] = [dc\cdot 2]X_3 + [dd\cdot 2]X_4 + \cdots,$$
u. s. w.

entsteht. Die Fortsetzung dieses Verfahrens führt schließlich zu einem reduzierten System, das nur eine Gleichung und nur die letzte Unbekannte enthält. Stellt man nun aus den Anfangsgleichungen sämtlicher Systeme das System

$$[na] = [aa]X_1 + [ab]X_2 + [ac]X_3 + \cdots,$$
$$[nb\,.\,1] = [bb\cdot 1]X_2 + [bc\cdot 1]X_3 + \cdots,$$
$$[nc\cdot 2] = [cc\cdot 2]X_3 + \cdots,$$
u. s. w.

zusammen, so läßt sich daraus, indem man schrittweise von unten herauf die Unbekannten bestimmt, ohne Schwierigkeit die gesuchte Lösung berechnen.

Fragt man jetzt, wie man dazu gekommen sei, gerade das Minimum von $[DD]$ als Bedingung einzuführen, so ist folgendes zu bemerken. Wenn, wie zunächst angenommen werden soll, die Größen n in den Widerspruchsgleichungen auf Messungen beruhen, die mit Fehlern zufälliger Art behaftet sind, und wenn ferner die Widersprüche zwischen den vorgelegten Gleichungen wesentlich von diesen zufälligen Beobachtungsfehlern herrühren, so ergeben sich aus den Erfahrungen, die über das Verhalten solcher Fehler bekannt sind, bestimmte Anhaltspunkte, die man selbstverständlich bei der Behandlung des vorliegenden Problems zu berücksichtigen hat. Diese Anhaltspunkte gestatten nun den Spielraum, der für die Lösungen des Systems (3) vorerst besteht, derart einzuengen, daß schließlich die Methode der kleinsten Quadrate zwar nicht als die einzige logisch zulässige, wohl aber als die einzige praktisch brauchbare Lösung übrig bleibt. Es ist hier nicht der Ort, den Nachweis für die soeben ausgesprochenen Sätze zu erbringen, weil dazu eine Entwicklung der Theorie der Beobachtungsfehler nötig wäre; es genügt darauf hinzuweisen, daß die oben erwähnte zusätzliche Bestimmung, welche zum Heraussuchen der annehmbarsten Lösung zu dienen hat, in diesem Falle auf den Erfahrungsunterlagen der Fehlertheorie beruht.

Anders stellt sich die Sache, wenn die Entstehung der Widersprüche nichts mit zufälligen Beobachtungsfehlern zu tun hat, sondern einfach davon herrührt, daß die darzustellende Funktion $f(x)$ mit der

darstellenden Funktion $F(x)$ nicht identisch ist, daß also eine Übereinstimmung zwischen beiden nur in einzelnen Punkten, aber nicht durchweg hergestellt werden kann. In solchen Fällen läßt sich die Benutzung der Methode der kleinsten Quadrate nur durch Zweckmäßigkeitsgründe motivieren. So kann man zunächst darauf hinweisen, daß das Minimum von $[DD]$ von vornherein alle diejenigen Lösungen ausschließt, deren Unbrauchbarkeit dadurch bedingt ist, daß man bei ihnen durch Abänderung der Unbekannten sämtliche Widersprüche gleichzeitig numerisch verkleinern kann: wenn das bei der Lösung der Methode der kleinsten Quadrate möglich wäre, so hätte man eben nicht das Minimum von $[DD]$ vor sich. Weiter läßt sich hervorheben, daß das Verfahren auf eine durchaus bestimmte Lösung und auf glatte Rechnungsvorschriften führt, deren Ziffernverbrauch sich nach Ausweis der Erfahrung in annehmbaren Grenzen hält. Diese Umstände schließen natürlich nicht aus, daß auch noch andere Lösungen der Aufgabe als brauchbar zuzulassen sind.

§ 108. Um das *Cauchy*sche Verfahren, das jetzt besprochen werden soll, zu erhalten, seien folgende Bemerkungen vorausgeschickt. Die m Gleichungen des Systems (3) sind so, wie sie dastehen, nicht zu verwenden, weil sie einander widersprechen. Auf der anderen Seite genügen p Gleichungen um ein Wertsystem für die p Unbekannten zu finden. Will man nun nicht etwa p Gleichungen willkürlich herausgreifen, so liegt es nahe, die p „Normalgleichungen“, die schließlich zur Bestimmung der Unbekannten dienen sollen, dadurch herzuleiten, daß man aus der Gesamtheit der vorgelegten m Gleichungen irgendwie p passende Kombinationen herstellt. Diese Kombinationen wird man natürlich linear wählen, um den rechnerischen Vorteil zu wahren, der durch die lineare Beschaffenheit des Systems (3) gegeben ist. Setzt man also ein System von Multiplikatoren $A_1, A_2, \cdots$ an und verbindet durch sie die Gleichungen (3), so erscheint eine erste Normalgleichung in der Gestalt

$$[nA] = [aA]X_1 + [bA]X_2 + \cdots,$$

die offenbar mit der Bedingung $[DA] = 0$ gleichbedeutend ist. In derselben Weise kann man mit einem zweiten System von Multiplikatoren $B_1, B_2, \cdots$ die zweite Normalgleichung in der Gestalt

$$[nB] = [aB]X_1 + [bB]X_2 + \cdots$$

bilden und auf diese Art fortfahren, bis man die p verlangten Normalgleichungen beisammen hat, die dann noch nach den Unbekannten aufzulösen sind. Nimmt man z. B. als Multiplikatoren $A, B, \cdots$ die Koeffizienten $a, b, \cdots$ selber, so kommt man offenbar wieder auf die Methode der kleinsten Quadrate.

Das vorstehende Verfahren soll jetzt etwas abgeändert werden. Zunächst werde mit den Multiplikatoren A wie angegeben die erste Normalgleichung gebildet, diese aber sogleich dazu benutzt, um die Unbekannte X_1 aus den einzelnen Widerspruchsgleichungen herauszuschaffen, sodaß die Widersprüche nunmehr in der von X_1 freien reduzierten Gestalt

$$D_k = n_{k1} - b_{k1} X_2 - c_{k1} X_3 - \cdot \cdot \qquad (6)$$

erscheinen. Daraus wird dann mit einem zweiten Multiplikatorensystem $B_1, B_2, \cdot \cdot$ die zweite, jetzt von X_1 freie, Normalgleichung hergeleitet, die wiederum sogleich dazu dient, die Unbekannte X_2 aus dem System (6) herauszuschaffen, sodaß ein drittes System von der Gestalt

$$D_k = n_{k2} - c_{k2} X_3 - d_{k2} X_4 - \cdot \cdot$$

entseht. In dieser Weise geht es weiter, bis zuletzt eine Normalgleichung mit der einen Unbekannten X_p entstanden ist. Wird diese benutzt, um aus dem letzten reduzierten System X_p zu eliminieren, so erscheinen die Widerspruchsgleichungen in der Gestalt

$$D_k = n_{kp},$$

d. h. die n_{kp} sind die Widersprüche, welche die aus den aufgestellten Normalgleichungen folgende Lösung übrig läßt.

Der hier beschriebene Rechnungsgang, der auch bei der Methode der kleinsten Quadrate anwendbar ist, besitzt den Vorteil, daß die Normalgleichungen sogleich in einer die Auflösung abkürzenden Gestalt erscheinen. Dafür ist allerdings für die jedesmal auszuführenden Eliminationen eine gewisse Mehrarbeit aufzuwenden, die jedoch unter Umständen durchaus angebracht sein kann. Will man nämlich wissen, wie sich der Anschluß jedesmal gestaltet, wenn man aus der Reihe der Unbekannten $X_1, X_2, \cdot \cdot$ nur eine oder nur zwei oder nur drei u. s. w. berücksichtigt und die folgenden von vornherein gleich Null setzt, so wird diese Frage sofort durch die Glieder $n_{k1}, n_{k2}, \cdot \cdot$ in den reduzierten Widerspruchgleichungen beantwortet.

§ 109. Die Einführung der Multiplikatorensysteme bringt die Gesamtheit der verschiedenen denkbaren Lösungen auf ein gemeinsames Schema, das auf die Ansetzung der Bedingungen

$$[DA] = 0, \quad [DB] = 0, \quad [DC] = 0, \quad \text{u. s. w.}$$

hinauskommt, die ihrerseits zu den Normalgleichungen führen. Handelt es sich nun nicht um die Ausgleichung von zufälligen Beobachtungsfehlern, sondern z. B. um den Anschluß von Kurve an Kurve, so ist die Aufgabe im allgemeinen ihrer Natur nach unbestimmt; im besondern verlieren dann auch die Gründe, die sich für die Methode der

kleinsten Quadrate geltend machen lassen, erheblich an Gewicht. Unter solchen Umständen liegt es nahe, von der Forderung auszugehen, daß sich die Rechnung möglichst bequem gestalten solle. Das wird nun bei dem zuletzt beschriebenen Verfahren erreicht, wenn man die Multiplikatoren $A, B, \cdots$ nur aus der Reihe der drei Zahlen $+1, -1, 0$ entnimmt, da in diesem Falle bei der Aufstellung der Normalgleichungen alle Multiplikationen fortfallen. Hierbei geht *Cauchy*, um die Sicherheit der Rechnung möglichst zu wahren, in folgender Weise vor.

Es bedeute das Symbol $\operatorname{sg}(x)$ wiederum wie in § 43 die Zahl $+1$ oder -1 oder 0, je nachdem x positiv oder negativ oder null ist. Man multipliziere jede Gleichung

$$D_k = n_k - a_k X_1 - b_k X_2 - \cdots$$

des vorgelegten Systems mit dem zugehörigen $\operatorname{sg}(a_k)$, sodaß in dem System der solchergestalt abgeänderten Gleichungen

$$D_k \operatorname{sg}(a_k) = n_k \operatorname{sg}(a_k) - a_k \operatorname{sg}(a_k) X_1 - \cdots \tag{7}$$

die Koeffizienten von X_1 durchweg negativ sind, abgesehen von dem Falle, wo ein a_k null ist. Dann soll die Summe des Systems (7) als die erste Normalgleichung dienen, die hiernach die Gestalt

$$0 = \sum_k D_k \operatorname{sg}(a_k) = N - A X_1 - B X_2 - \cdots$$

erscheint, wo zur Abkürzung

$$N = \sum_k n_k \operatorname{sg}(a_k), \quad A = \sum_k a_k \operatorname{sg}(a_k), \quad B = \sum_k b_k \operatorname{sg}(a_k), \cdots$$

gesetzt ist. Schafft man hiermit, wie angegeben, X_1 aus dem System (7) fort, so entsteht das erste reduzierte System in der Gestalt

$$D_k \operatorname{sg}(a_k) = n_{k1} - b_{k1} X_2 - c_{k1} X_3 - \cdots, \tag{8}$$

wo

$$n_{k1} = n_k \operatorname{sg}(a_k) - \frac{N a_k \operatorname{sg}(a_k)}{A}, \quad b_{k1} = b_k \operatorname{sg}(a_k) - \frac{B a_k \operatorname{sg}(a_k)}{A}, \tag{9}$$

u. s. w.

ist. Hieraus ergeben sich noch die Beziehungen

$$0 = \sum_k n_{k1} = \sum_k b_{k1} = \sum_k c_{k1} = \cdots,$$

die eine bequeme und nützliche Kontrolle für die Richtigkeit der ersten Normalgleichung und der Elimination enthalten.

§ 110. Der beschriebene Rechnungsgang leidet noch an der Unbequemlichkeit, daß in den Gleichungen mit negativem a erst alle Vorzeichen umzukehren sind. Diese Umkehrung läßt sich indessen durch einen einfachen Kunstgriff umgehen, wenn man das Rechen-

schema passend anlegt. Man setzt zunächst die Größen $n, a, b, \cdots$ spaltenweise nebeneinander, wobei die positiven und die negativen Zahlen ihre eigenen Spalten erhalten, die mit $n+, n-, \cdots$ bezeichnet werden mögen. Ferner läßt man zwischen je zwei benachbarten Gleichungen eine Zeile frei; die leeren Räume dienen nachher dazu, die Zwischengrößen der Eliminationsrechnung aufzunehmen. Um nun die Größen $N, A, \cdots$ für die erste Normalgleichung zu bilden, deckt man mit Karton- oder Blechstreifen von passender Breite zunächst die Gleichungen ab, deren a negativ oder null ist, und bildet die Spaltensummen der unbedeckten Zahlen, die mit $n++, n-+, \cdots$ bezeichnet werden mögen. Danach deckt man die Gleichungen ab, deren a positiv oder null ist, und summiert wiederum spaltenweise die unbedeckten Zahlen, wobei die Summen $n+-, n--, \cdots$ entstehen. Aus diesen Teilsummen ergeben sich nun, wie leicht zu sehen, die Größen $N, \cdots$ durch die Operation

$$(n++)-(n+-)-[(n-+)-(n--)], \quad \text{u. s. w.}$$

Nachdem die Normalgleichung $N = AX_1 + \cdots$ hergestellt ist, setzt man unter die Größen $n_k, b_k, c_k, \cdots$ die Werte der Ausdrücke

$$a_k N : A, \quad a_k B : A, \quad a_k C : A, \cdots$$

und subtrahiert; die entstehenden Differenzen, die — von dem Faktor $\operatorname{sg}(a_k)$ abgesehen — mit den in (9) angesetzten Größen übereinstimmen, werden auf den Zeilen der zugehörigen Gleichungen in die Spalten eingetragen, die man für die Größen des ersten reduzierten Systems vorbereitet hat, wobei jedoch die Gleichungen, in denen a_k null ist, vorläufig auszulassen sind. Für die neuen Spalten müssen dann, wenn man wieder mit den vorhin beschriebenen Abdeckungen summiert, die Kontrollgleichungen

$$(++)-(+-)-[(-+)-(--)] = 0$$

erfüllt sein. Ist die Kontrolle erledigt, so überträgt man noch, wenn nötig, die bisher nicht berücksichtigten Gleichungen, in denen a_k gleich Null ist, nach dem ersten reduzierten System und hat damit alles vorbereitet, was zur Bildung der zweiten Normalgleichung erforderlich ist. Die Rechnung erfolgt dann genau so wie vorher, nur daß sich jetzt die Abdeckungen nach den b-Größen richten.

Wenn die Anzahl der Widerspruchsgleichungen nicht groß ist, so kann man statt abzudecken die Multiplikation mit den sg-Faktoren in der Weise ausführen, daß man die Zahlen, welche den Faktor $\operatorname{sg} = -1$ erhalten sollen, ausstreicht und — am besten mit andersfarbiger Tinte — in der anderen Abteilung der zugehörigen Doppelspalte wieder einträgt, während die Zahlen, welche den Faktor $\operatorname{sg} = 0$

erhalten, einfach ausgestrichen und erst nach Bildung der betreffenden Normalgleichung in das nächste reduzierte System übertragen werden.

Das *Cauchy*sche Verfahren ist wegen der einfachen Bildungsweise der Normalgleichungen sowie wegen der beständigen Kontrollen sehr bequem auszuführen; gleichwohl begegnet man ihm in der Literatur nur sehr selten. Bei näherer Betrachtung kann dies auch nicht sonderlich überraschen: der Fall, daß die Widersprüche von zufälligen Fehlern herrühren und deshalb die Anwendung der Methode der kleinsten Quadrate verlangen, kommt unvergleichlich viel häufiger vor, als der Fall einer von Beobachtungsfehlern unabhängigen Interpolation: es ist daher leicht verständlich, daß der Rechner die Methode der kleinsten Quadrate, die ihm ohnehin geläufig geworden ist, schließlich auch da vorzieht, wo sie nicht unmittelbar geboten, sondern nur zulässig ist.